AD-A049 707 GEORGIA UNIV ATHENS DEPT OF STATISTICS AND COMPUTER--ETC F/G 12/1
INTERACTIVE STATISTICAL SOFTWARE.(U)
JAN 78 R E BARGMANN, H BOUVER N00014-69-A-0423
UNCLASSIFIED THEMIS-UGA-33 NL

1 OF 2
AD
A049707

(24) (12)

OFFICE OF NAVAL RESEARCH

CONTRACT N00014-69-A-0423, NR 042-261

FINAL REPORT

ROLF E. BARGMANN & HUBERT BOUVER
EDITORS

Rolf Bargmann
Principal Investigator

The University of Georgia
Department of Statistics and Computer Science
Athens, Georgia

January 1978

(See 1473)

Grant Period: April 1, 1969 to September 30, 1974

Associate Investigators

Dr. George Adomian	69/72
Dr. Hubert Bouver	71/74
Dr. Niels Eskelson	69/71
Dr. Herman Kopetz	69/71
Dr. Frank G. Lether	70/73
Dr. J. David Mason	70/72
Dr. James E. Norman	69/72
Dr. Takis Papaioannou	71/72
Dr. Lucius Penn	71/74

TITLE:

INTERACTIVE STATISTICAL SOFTWARE

Research Assistants

Ann A. Ballengee	70/72
Steven F. Bingham	72/74
Walter P. Bond	71/74
L. Edward Cannon	70/71
Chit Fu Chang	69/72
Carl R. Fortson	69/70
Richard W. Graney	69/70
Rebecca Grant	69/70
Judy L. Hayward	72/74
Judy L. Ishee	69/70
Lois J. Knybel	69/71
Kenneth R. Kundert	71/72
M. Elizabeth Nash	72/73
H. I. Patel	69/71
Robert I. Schwartz	69/70
Janice S. Scott	70/71
Carlton G. Thomas	69/70
S. J. Trivedi	69/71
W. Walker	70/71
R. L. Wood	72/73
Chang-Wu Yen	71/72

TABLE OF CONTENTS

APPENDICES

(Technical Documentation, Computer Programs, Displays)

Volume I

Volume II

Volume III

Volume IV

Volume V

Volume VI

INTRODUCTION

This is the narrative portion of Final Report of ONR Contract No. N0014-69-A-0423, NR 042-261. The final technical report (documentation, user manuals, displays, tables) was prepared and distributed in six volumes, entitled Final Report, Appendix Volumes I to VI, in July, 1974. AD-783893, 894, 895; 786475 + 476 786493 Since this narrative makes considerable use of these appendices, they should be regarded as an integral part of this Final Report.

This summary report is, to some extent, a reflection of the developments of interactive computation from 1969 to 1975. Very considerable systems programming work was required in the initial stages. In 1970 we were fortunate to receive a copy of the GRAF package from the Health Facility of UCLA. After some work, especially to fit these routines into our narrower scope, this tool enabled us to obtain operational modules in 1971 and 1972. By that time we had developed our own monitor system, better adapted to our needs, especially for statistical systems. A manipulative calculator unit involving statistical distribution routines (central and non-central) became available and was integrated with several of the later units (response surfaces, interactive OMNITAB). From the view of 1977 this may appear fairly common, but in 1971 there were very few systems having that facility.

The rapid changes of computer configurations, and availability of cheaper terminals that could handle all but the most elaborate displays of the IBM 2250 graphics system, made our system somewhat anachronistic by 1973, and efforts were begun to adapt or prepare interactive units for other configurations. However, in the context of this project we resisted from preparing "statistical packages" and, where we needed no graphics, presented statistical routines as library functions, easily accessible in FORTRAN. Since 1974, this part of our interactive software (see, especially, Bouver and Bargmann [42]) was used much more frequently than the graphics terminal.

In 1974 we took the 2840 Control Unit off maintenance, because the teletype terminals widely available then had become less expensive to purchase than a one-year maintenance agreement on the graphics system. Some of the adaptation work was done in 1973, and is described in this report (Hayward and Bargmann [38]); this process is still in progress. As units are being adapted and documented, credit will be given to this grant, in which the original design work was made.

The changes in hardware (360 to 370, changes of disk drives and core configurations) and environment (variable storage(MVS)), which took place between 1973 and 1976, required considerable reprogramming, especially since procedures in the newer configurations were not readily adaptable to the older (1965) hardware and channel configuration. There were delays, often extending over several months, but we did succeed in getting the interactive graphics system operational under the latest configuration, in October/November of 1977. Except where otherwise noted, the photographs of the video-screen exhibited in this report were made from the latest version.

As anticipated in the original THEMIS proposal of 1969, the major part of the software development was done by graduate students, who performed these tasks in conjunction with their thesis or dissertation work. In fact, three of the major systems and applications units constituted Ph.D. dissertations devoted almost entirely to the development of statistical software; ("An On-Line Statistical System for Lay Usage", Penn [12]; "Interactive OMNITAB for Statistical Usage", Bingham and Bargmann [41]; and "Graphical Aids for Statistical Computation", Bond and Bargmann [43]). Except for the very earliest tasks (the prior THEMIS task had a broader spectrum of coverage) all research reports, theses, and dissertations prepared under this grant included substantial work in interactive statistical computation, to which the task was liminted since 1970.

Transportability has never been a problem for the mathematical and statistical subroutines, since all were written in FORTRAN. The major difference in the programs for CDC Cyber and IBM is the need for double precision in the latter, which is seldom required in the former. For the device-specific graphics subroutines the problem

was quite difficult. An adaptation was successfully made at the University of Arkansas, in 1972. The routines compiled with the aid of the UCLA GRAF system are not easily reconstructed from the source programs. The load modules need to be changed, directly (ZAP), to accommodate changes in configuration. The version presently on tape has been shown to work successfully on an IBM 370/158 (MVS) via a 2840 mod I controller and an IBM 2250 console (which need not have absolute vector capability); see Chapter 3.

The graphics monitor system has had considerable use in research and classroom work. At one time (1972-73) it was available 24 hours a day and often used for preparation of demonstrations, and in thesis research. The batch versions of some of the larger units have even wider use inside and outside the University of Georgia. Since the availability of "express" runs and quick turnaround even of batch jobs, some of the units requiring large output (e.g., Analysis of Covariance, Hierarchical multivariate analysis) can be handled promptly in batch mode, even though repeated user interference is required.

This narrative report consists of five chapters: Chapters 1 and 2 describe interactive graphical units and present examples of the use of these units, with photographs of the screen taken in the Fall of 1977, when the intergrated system was operational on the latest monitor. Chapter 3 consists of a short description of the graphics system, including examples and considerations of transportation and adaptation to other systems. For the detailed description of these programs and their uses the reader is referred to the technical documentation in the appendix volumes or THEMIS reports.

Chapter 4 describes those tasks which supported the software development, especially numerical analysis work and the development and testing of efficient and precise modules. Other statistical tasks performed under this grant are described in Chapter 5.

Technical details, instructions to users and programmers, important formulas, and tables and selected displays of program uses are contained in six appendix volumes to this Final Report, which were

distributed in the Fall of 1974, and to which frequent reference has been made in this narrative portion.

CHAPTER 1

First Phase: Programs under the UCLA GRAF MONITOR

1.1 Interactive Input-Output Analysis

Documentation: Appendix C, appendix Vol. I, pp 50-132, and THEMIS Report No. 11, Fortson [15].

After a call to the program segment $LINK MODEL the user is required to describe his network in terms of flow of materials from Input to Intermediate to Output products, and efficiencies (Leontieff algorithms). A forecaster will use this network by calling $LINK FORECAST and answering questions on demand of output, demand (if any) of intermediate products, variation of demand, and dependence of demand on general economic trend (a one-factor factor analytic model is used for the estimation of the covariance matrices). Changes in product flow can be made, at this stage, with a light pen. Novel features in this unit are:

> Establishment of a variance-covariance matrix of demand, and effect on the variance of input requirements, in the Leontieff model.

The user is expected to express, as a coefficient of determination (in percent) the relationship between sales of a product and general economic activity. This relationship may be negative (e.g., for spare parts). The program constructs a one-factor factor analytical model for the variance-covariance matrix of demands. In a class taught at the Lockheed program of the Georgia Institute of Technology (1971), some students investigated, by simulation, the effect on requirements if the underlying model has, in fact, two "common factors." Changes were quite insignificant.

The final demonstration was made by Mr. Satish Mehra (in conjunction with his dissertation work in Management) in October, 1977, to study the sensitivity of requirements to estimates of the dependence of demand on general economic trends. Sample displays were photographed from the IBM 2250 screen and are shown in Figures 1.1.1 to 1.1.8. The data were based on an actual production and inventory network. The study showed that a wide range of trial values for the coefficient of variation (from zero to 20 percent to the empirically found 53 percent)

```
PRODUCTS 1 THRU     9 REFER TO INPUTS
PRODUCTS    10 THRU    14 REFER TO INTERMEDIATES
PRODUCTS    15 THRU    18 REFER TO OUTPUTS


 THIS SECTION OF THE INPUT-OUTPUT ANALYSIS ALLOWS YOU TO
MAKE SALES FORECASTS AND TO PLACE UPPER AND LOWER BOUNDS ON THE
AMOUNT OF RAW MATERIAL (INPUT PRODUCTS) NECESSARY
WE WILL ALSO CONSIDER THE VARIATION ON THE DEMAND OF THE OUTPUT
PRODUCTS AND THE INFLUENCE OF A COMMON FACTOR (SUCH AS THE GROSS
NATIONAL PRODUCT) ON THE SALE OF EACH PRODUCT

 AFTER YOUR RESULTS ARE DISPLAYED YOU WILL BE GIVEN THE OPPORTUNITY
TO CHANGE VARIOUS PARAMETERS AND RECALCULATE THE RESULTS

 WE WILL NOW ESTABLISH THE DEMANDS FOR THE PRODUCTS
YOU WILL BE ASKED TO SUPPLY THE AMOUNT OF EACH OUTPUT PRODUCT
DESIRED AND THE AMOUNT OF ANY INTERMEDIATE PRODUCTS TO BE PURCHASED
OR SOLD

 PRESS KEY 1 IF ANY INTERMEDIATE PRODUCTS ARE TO BE PURCHASED
PRESS KEY 2 IF NO INTERMEDIATE ARE TO BE PURCHASED
PRESS KEY 3 TO MODIFY MODEL BEFORE PLACING DEMANDS
```


Fig. 1.1.1

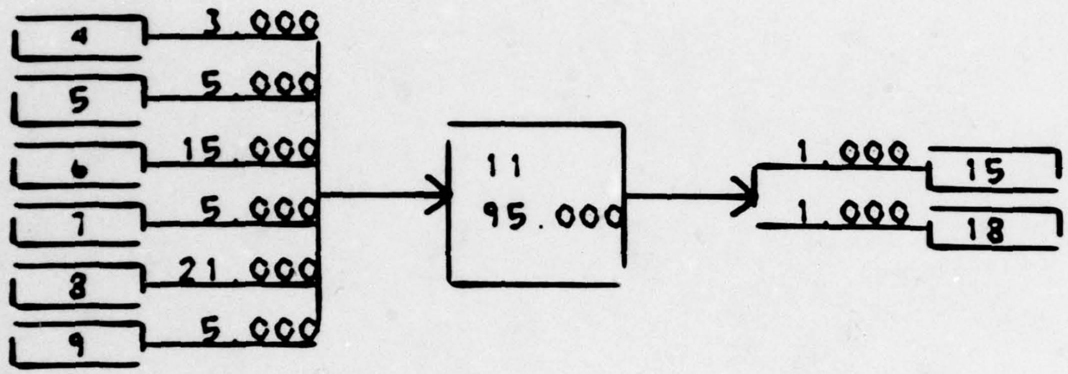

4
5
6
7
8
9
3.000
5.000
15.000
5.000
21.000
5.000
11
95.000
1.000
1.000
15
18

Fig. 1.1.2

```
.ODUCTS 1 THRU      9 REFER TO INPUTS
.ODUCTS    10 THRU     14 REFER TO INTERMEDIATES
RODUCTS    15 THRU     18 REFER TO OUTPUTS


HOW MANY INTERMEDIATE PRODUCTS WILL BE SOLD?3.

PLEASE FILL IN THE FOLLOWING BLANKS

120    UNITS OF PRODUCT 12.     WILL BE SOLD
80     UNITS OF PRODUCT 13.     WILL BE SOLD
100    UNITS OF PRODUCT 14.     WILL BE SOLD

DATA ON SCREEN IS TECHNICALLY CORRECT
PLEASE RECHECK AND PRESS KEY 1 IF YOU AGREE
PRESS KEY 2 TO REENTER DATA
```

Fig. 1.1.3

```
RODUCTS 1 THRU      9 REFER TO INPUTS
RODUCTS    10 THRU    14 REFER TO INTERMEDIATES
RODUCTS    15 THRU    18 REFER TO OUTPUTS

WE WILL NOW ESTABLISH THE DEMANDS ON THE OUTPUT PRODUCTS

PLEASE FILL IN THE FOLLOWING BLANKS

300    UNITS OF PRODUCT    15 WILL BE SOLD
400    UNITS OF PRODUCT    16 WILL BE SOLD
400    UNITS OF PRODUCT    17 WILL BE SOLD
300    UNITS OF PRODUCT    18 WILL BE SOLD

DATA ON SCREEN IS TECHNICALLY CORRECT
PLEASE RECHECK AND PRESS KEY 1 IF YOU AGREE
PRESS KEY 2 TO REENTER DATA
```

Fig. 1.1.4

```
PRODUCTS 1 THRU      9 REFER TO INPUTS
PRODUCTS   10 THRU    14 REFER TO INTERMEDIATES
PRODUCTS   15 THRU    18 REFER TO OUTPUTS


 WE WILL NOW ESTABLISH VARIATIONS ON THE DEMANDS OF THE PRODUCTS
THAT WILL BE SOLD

 YOU WILL NEED TO PLACE A LIMIT ON THE AMOUNT OF VARIATION THAT
MAY BE EXPECTED ON THE SALE OF EACH PRODUCT
EXAMPLE  IF YOU EXPECT TO SELL 40 UNITS OF PRODUCT 8 PLUS
OR MINUS 10 UNITS: THE VARIATION WOULD BE 10. YOU MAY ALSO
KNOW THIS AS THE COMMON '3-SIGMA' LIMIT

 WE WILL ALSO NEED TO ESTABLISH THE RELATIONSHIP OF THE VARIATION BETWEEN
PRODUCTS  WE WILL DO THIS BY RELATING EACH PRODUCT TO THE OVERALL
ECONOMY  THIS COEFFICIENT OF DETERMINATION MUST BE BETWEEN +100 AND -100
THE MINUS VALUES INDICATE THAT SALES ON THE PRODUCT DROP AS THE
OVERALL ECONOMY INCREASES
AS A GUIDELINE I SUGGEST A COEFFICIENT OF DETERMINATION OF THE ORDER OF 15

FOR NECESSITIES SUCH AS FOOD AND VALUES SUCH AS 10 FOR NON-ESSENTIALS
SUCH AS COLOR TV SETS
REPAIR PARTS MIGHT HAVE A VALUE SUCH AS -10

 PLEASE FILL IN THE FOLLOWING BLANKS

THE VARIATION OF PRODUCT    12 IS _       UNITS
```

Fig. 1.1.5

```
PRODUCTS 1 THRU      9 REFER TO INPUTS
PRODUCTS   10 THRU    14 REFER TO INTERMEDIATES
PRODUCTS   15 THRU    18 REFER TO OUTPUTS


 PLEASE FILL IN THE FOLLOWING BLANKS

THE VARIATION OF PRODUCT   12 IS 90.    UNITS
THE VARIATION OF PRODUCT   13 IS 90.    UNITS
THE VARIATION OF PRODUCT   14 IS 150.   UNITS
THE VARIATION OF PRODUCT   15 IS 180.   UNITS
THE VARIATION OF PRODUCT   16 IS 180.   UNITS
THE VARIATION OF PRODUCT   17 IS 120.   UNITS
THE VARIATION OF PRODUCT   18 IS 210.   UNITS

 DATA ON SCREEN IS TECHNICALLY CORRECT: PLEASE RECHECK
AND PRESS KEY 1 IF YOU AGREE
PRESS KEY 2 TO REENTER DATA
```

Fig. 1.1.6

```
ODUCTS 1 THRU      9 REFER TO INPUTS
ODUCTS    10 THRU     14 REFER TO INTERMEDIATES
RODUCTS    15 THRU     18 REFER TO OUTPUTS

THE COEFFICIENT OF DETERMINATION OF PRODUCT   12 IS -20
THE COEFFICIENT OF DETERMINATION OF PRODUCT   13 IS 0
THE COEFFICIENT OF DETERMINATION OF PRODUCT   14 IS -10
THE COEFFICIENT OF DETERMINATION OF PRODUCT   15 IS 51.
THE COEFFICIENT OF DETERMINATION OF PRODUCT   16 IS 51.
THE COEFFICIENT OF DETERMINATION OF PRODUCT   17 IS 51.
THE COEFFICIENT OF DETERMINATION OF PRODUCT   18 IS 20.

DATA ON SCREEN IS TECHNICALLY CORRECT: PLEASE RECHECK
AND PRESS KEY 1 IF YOU AGREE
PRESS KEY 2 TO REENTER DATA
```

Fig. 1.1.7

REQUIREMENTS OF INPUT PRODUCTS

PRODUCT NO	LOWER LIMIT	UPPER LIMIT
1	367.37	1106.32
2	611.12	1704.67
3	1538.22	4567.04
4	6864.25	15628.80
5	2912.14	10277.87
6	8745.76	33359.46
7	3282.19	12230.26
8	35022.28	86351.56
9	2403.59	8444.05

PRESS KEY 1 TO CHANGE DEMAND, VARIATION, OR CONFIDENCE LIMITS
PRESS KEY 2 TO CHANGE MODEL
PRESS KEY 31 TO TERMINATE

Fig. 1.1.8

produced only very small fluctuations.

This unit was demonstrated at N. C. State College in Raleigh, N. C. in 1970, at the Iowa State University in Ames, Iowa, in 1972, and at several national meetings of the American Statistical Association. It was used in our Information Systems courses (STA 804). Many visiting groups saw this unit in operation. It was implemented at the University of Arkansas in 1973.

An adaptation to non-graphics terminal use is now under development.

1.2 Interactive Quantal Analysis

Documentation: Appendix H, Appendix Vol II, pp. 246-281, and THEMIS Report No. 16 (Ishee [21]).

To perform this analysis commonly known as Bio-Assay the user connects to $LINK QUANTAL and fills in the blanks shown in Figure 1.2.1. After data entry is complete the user has the option to ask for analysis by Probit, Logit, log-log, arcsine, or Weibull transformation with shape parameters 1 to 6. Figure 1.2.2 appears on the screen and indicates how well the data are approximated by the chosen growth curve.

Figures 1.2.3 and 1.2.4, are examples of numerical results. Figure 1.2.5 indicates whether the selected transformation was chosen properly, At this stage the user has a variety of options to transform dosages (e.g. logs, square roots), and to pre-view figures such as 1.2.5 before re-doing the analysis with the new transformations.

> Novel features: Really standard quantal analysis, but with preview option to save unnecessary computer time. Also the inclusion of the Weibull transformation of degree 3 or greater is non-standard. The ease with which best transformations can be found, by combination of batches and scale changes, was the reason for the popularity of this unit.

Until 1974, when "Interactive OMNITAB" became operational, the QUANTAL unit was by far the most frequently used module. It was used in studies of suicide trends, design of atomizer nozzles in insecticide spray, studies on efficiency of family planning clinics, and many others. A very simple interactive version exists (without graphics) on our CDC Cyber computer. The algorithms are also used in the WEIBUL module (see Section 2.1. below) and are as appropriate in the fitting of growth functions as are order statistics and, of course, much faster.

Demonstrations, with transparencies, were shown at several national meetings in 1971-73. They were also shown to staff members at Fort Benning, Ga., and used for data analysis of small arms simulation studies. The unit was operational at the University of Arkansas, in 1973.

```
     NEXT COMPLETE THE FOLLOWING STATEMENT FOR EACH OF THE     8
BATCHES.  AN EXAMPLE IS, IN BATCH 12. WHERE THE CONCOMITANT VARIABLE
(DOSAGE)     6.4        WAS GIVEN TO 36. SUBJECTS, THERE WERE 30. SUBJECTS
THAT RESPONDED (SUCCESSES).

     THE NUMBER OF SUBJECTS IN EACH BATCH MUST BE GREATER THAN 2.

     IN BATCH      1 WHERE THE CONCOMITANT VARIABLE (DOSAGE) 1.
WAS GIVEN TO 50.   SUBJECTS, THERE WERE  2.     SUBJECTS THAT RESPONDED
(SUCCESSES).

     IN BATCH      2 WHERE THE CONCOMITANT VARIABLE (DOSAGE) 1.5
WAS GIVEN TO 50.   SUBJECTS, THERE WERE  5.     SUBJECTS THAT RESPONDED
(SUCCESSES).

     IN BATCH      3 WHERE THE CONCOMITANT VARIABLE (DOSAGE) 2.
WAS GIVEN TO 40.   SUBJECTS, THERE WERE  7.     SUBJECTS THAT RESPONDED
(SUCCESSES).

     IN BATCH      4 WHERE THE CONCOMITANT VARIABLE (DOSAGE) 3.
WAS GIVEN TO 30.   SUBJECTS, THERE WERE  10.    SUBJECTS THAT RESPONDED
(SUCCESSES).

     IN BATCH      5 WHERE THE CONCOMITANT VARIABLE (DOSAGE) 3.5
WAS GIVEN TO 20.   SUBJECTS, THERE WERE  12.    SUBJECTS THAT RESPONDED
(SUCCESSES).

     IN BATCH      6 WHERE THE CONCOMITANT VARIABLE (DOSAGE) 4.
WAS GIVEN TO 40.   SUBJECTS, THERE WERE  30.    SUBJECTS THAT RESPONDED
(SUCCESSES).

     IN BATCH      7 WHERE THE CONCOMITANT VARIABLE (DOSAGE) 4.5
WAS GIVEN TO 30.   SUBJECTS, THERE WERE  26.    SUBJECTS THAT RESPONDED
(SUCCESSES).

     IN BATCH      8 WHERE THE CONCOMITANT VARIABLE (DOSAGE) 5.5
WAS GIVEN TO 50.   SUBJECTS, THERE WERE  48.    SUBJECTS THAT RESPONDED
(SUCCESSES).
```

Fig. 1.2.1

THE GRAPH BELOW DISPLAYS THE OBSERVED AND PREDICTED PROPORTION OF SUCCESS FOR EACH BATCH (CLASS).

THE HORIZONTAL AXIS IS THE X-DOSAGE OF EACH BATCH (ARRANGED IN ASCENDING ORDER). THE BATCH NUMBER FOR EACH CORRESPONDING BATCH IS DIRECTLY UNDER THE AXIS AND THE TWO LOWER FIGURES ARE THE MINIMUM AND MAXIMUM X-DOSAGE.

THE VERTICAL AXIS IS THE PROPORTION OF SUCCESSES AXIS. THE LINED GRAPH IS THE PREDICTED PROPORTION OF SUCCESSES AND THE ASTERICKS DENOTE THE OBSERVED PROPORTION OF SUCCESSES.

WHEN YOU ARE READY TO PROCEED, DEPRESS PROGRAM FUNCTION KEY 1.

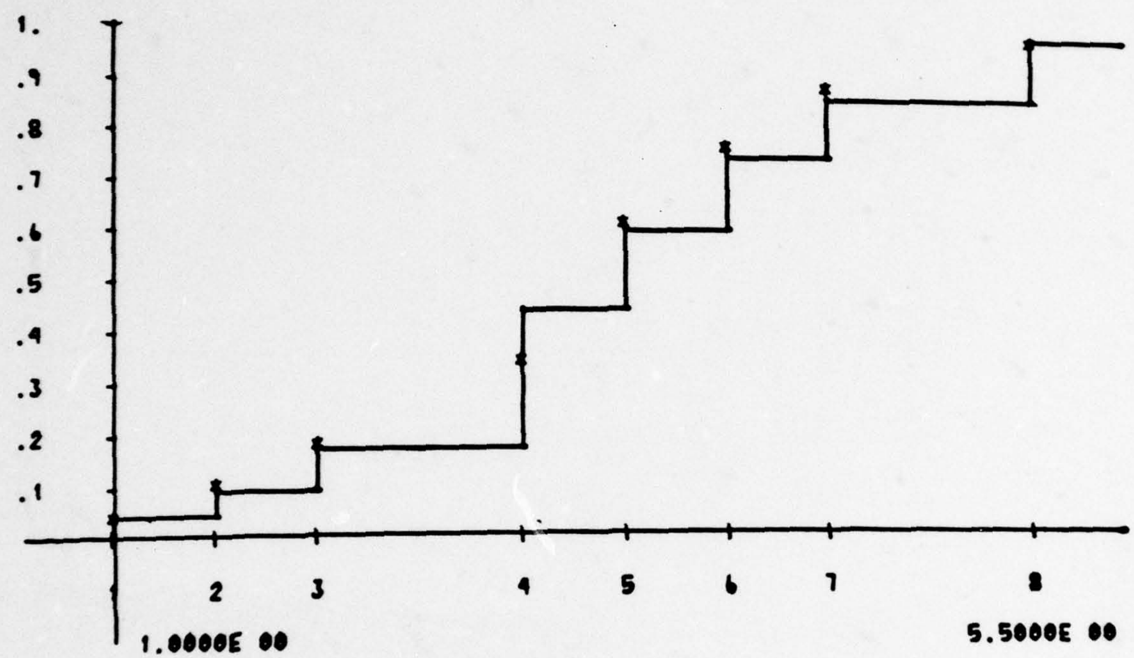


Fig. 1.2.2

LOGIT ANALYSIS

THIS IS THE LAST PAGE OF THIS SECTION OF YOUR OUTPUT. PAGE 1

IF YOU WISH TO REVIEW THIS SECTION, DEPRESS PROGRAM FUNCTION KEY 4

OTHERWISE, DEPRESS PROGRAM FUNCTION KEY 1.

CLASS	1	2	3	4	5
SIZE	50	50	40	30	20
P(OBS)	4.0000E-02	1.0000E-01	1.7500E-01	3.3333E-01	6.0000E-01
P(PRED)	4.4293E-02	8.4340E-02	1.5473E-01	4.1963E-01	5.8965E-01
Y(PRED)	-3.0716E 00	-2.3848E 00	-1.6980E 00	-3.2430E-01	3.6253E-01
Y(OBS)	-3.1781E 00	-2.1972E 00	-1.5506E 00	-6.9315E-01	4.0546E-01
Y(WORK)	-3.1730E 00	-2.1820E 00	-1.5430E 00	-6.7863E-01	4.0529E-01
WEIGHTS	2.1166E 00	3.8613E 00	5.2316E 00	7.3062E 00	4.8392E 00
XDOSAGE	2.7183E 00	4.4817E 00	7.3891E 00	2.0086E 01	3.3115E 01
Y-ERROR	-1.0643E-01	1.8757E-01	1.4736E-01	-3.6885E-01	4.2931E-02

CLASS	6	7	8
SIZE	40	30	50
P(OBS)	7.5000E-01	8.6667E-01	9.6000E-01
P(PRED)	7.4065E-01	8.5020E-01	9.5730E-01
Y(PRED)	1.0494E 00	1.7362E 00	3.1099E 00
Y(OBS)	1.0986E 00	1.8718E 00	3.1781E 00
Y(WORK)	1.0980E 00	1.8655E 00	3.1760E 00
WEIGHTS	7.6834E 00	3.8207E 00	2.0439E 00
XDOSAGE	5.4598E 01	9.0017E 01	2.4469E 02
Y-ERROR	4.9247E-02	1.3560E-01	6.8193E-02

Fig. 1.2.3

LOGIT ANALYSIS

ANALYSIS OF VARIANCE TABLE

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE
REGRESSION	0.97694717E 02	1	0.97694717E 02
ERROR	0.13233643E 01	6	0.22056067E 00
TOTAL	0.99018082E 02	7	

CALCULATED F = 0.44293799E 03

IF YOU WISH TO LEAVE THE QUANTAL PROGRAM TO DETERMINE F(1, 6),BY CALLING THE "CALCG" PROGRAM, DEPRESS PROGRAM FUNCTION KEY 11. IF YOU WISH TO CONTINUE, DEPRESS PROGRAM FUNCTION KEY 1.

Fig. 1.2.4

THE FOLLOWING IS THE GRAPH OF THE STRAIGHT LINE Y = ALPHA + BETA * X
WHERE ALPHA AND BETA HAVE BEEN ESTIMATED IN THE PROBIT ANALYSIS.
HERE THE EQUATION IS Y = -0.254189E 01 + 0.788234E 00 * X.
THE HORIZONTAL AXIS IS THE DOSAGE-AXIS AND THE VERTICAL AXIS IS THE Y-AXIS

THE VALUES DENOTED BY '*' ARE THE OBSERVED Y-VALUES

WHEN YOU ARE READY TO PROCEED, DEPRESS PROGRAM FUNCTION KEY 1.

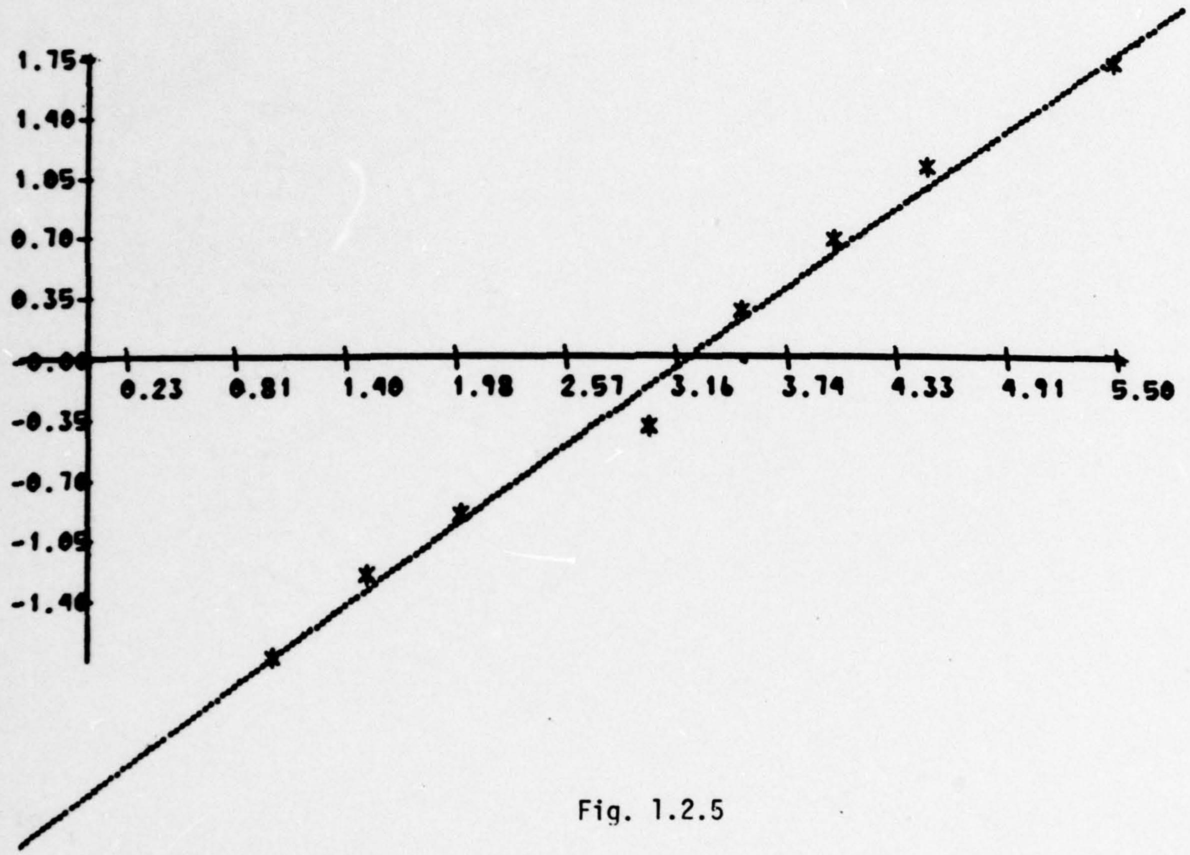


Fig. 1.2.5

This unit was fully operational at the last test, in October, 1977. However, the photographs are from an earlier run, since the quality of the more recent photographs is inferior.

1.3. Cluster Configurations in n Dimensions.

Documentation Appendix I, appendix vol. III, pp. 1-71 and Themis Report No. 17 (Trivedi [23])

In our recent test under MVS we were unable to obtain photographs of this unit. The reader is advised to look at the displays in Appendix Vol. III and the instructions in Themis Report No. 17, pp. 68-113.

The development of this interactive graphical unit, which requires batch pre-processing of data, was prompted by the desire to have a combined factor-analysis and cluster-analysis algorithm which could be used when the strongest factors or clusters are trivial. The dilemma is best seen in an example which I ran at IBM many years ago. The factor analysis of symptoms of diseases was overshadowed by one factor, consisting of the trivial syndrome fever-pain-and-chills, and the cluster analysis of patients identified just clusters of very ill, slightly ill, and healthy patients, without differerentiation of diseases. A simple structure rotation within clusters was attempted, to see whether the patients fell into an overdetermined subspace only in the clusters of ill people, and not in others. Many workers in multivariate analysis would like to "see" their configurations as projections of p-dimensional hyper-ellipsoids into ellipses in 2-dimensional subspaces.

The data which were generated were supposed to exhibit such secondary structure - alignment into overdetermined subspaces ("simple structure planes") in only some of the clusters. When we knew the location of these subspaces, and chose projection accordingly, they were revealed at once (page 86 and 88 of Themis Report No. 17, page 25 of

Appendix Volume III). In much the same way as factor analysts did their rotations by graphical methods, for all pairs of factors, before the advent of computers, we developed this module to display such projections instantly on a screen - not just pairs of orthogonal axes but any projection on which we might suspect to find significant overdetermination.

When one looks at one projection of a four-dimensional ellipsoid in 2-space (see page 104 of Themis Report No. 17) one really has no inkling of the underlying configuration when viewed from a different projection (see page 105).

There has been much experimentation with this unit, and it was displayed and discussed in several symposia (here at the University of Georgia, and at the North Carolina State University at Raleigh), but it appears that identification of overdetermined subspaces in n-dimensions is still as much of an art as it has always been.

A very elaborate pre-processing program, described on pages 68-71 of Themis Report No. 17, gives the user several pointers where to look for overdetermined subspaces. The display is then very simple; the user merely calls $LINK ELLIPSE and follows the simple instructions (the reader is referred to pages 7-25 of Final Report, Appendix Vol. III). The unit has been used in classroom instruction in a course of multivariate analysis. The batch pre-processor has been used, by itself, for several studies, especially in Food Science.

We did not succeed in making this unit operational on the 370/158 under MVS - it worked well under earlier releases and at the University of Arkansas. It is expected that changes will have to be made on the device-dependent segments (e.g., last buffer address) which need to be set differently in different installations. The corresponding code is in the section relating to the GRAF (UCLA) monitor.

1.4 <u>Other Units Under UCLA GRAF Program</u>.

<u>Queuing Analysis for Branched Processes</u>

Documentation: Appendix D, appendix Vol. I, pp. 132 to 178, and THEMIS Report No. 12 (Knybel [16]).

This was an interactive program to obtain distributions of

queues (typically in a manufacturing process) under an M/M/s queuing discipline (exponential arrival, exponential departure, s servers in a channel). This unit was used in classroom instruction prior to 1974, when it was replaced by the more versatile PRODFLOW unit (see Section 2.5).

A Conversational Unit for Hierarchical Discriminant Analysis

Documentation: Appendix E, Appendix Vol. I, pp. 178-227, and THEMIS Report No. 13 (Schwartz [17]).

First attempt to obtain an interactive version of a multivariate analysis program for irregular data, with plotting facilities for a 2-way analysis of variance. Was replaced one year later by the more general "Interactive Multivariate Data Analysis Program" (see Section 2.4).

CHAPTER 2

Second Phase: Program under the University of Georgia Monitor System (GMS, COMAP, COMFORT)

For narrative description of the system, adaptation, and transportation, see Chapter 3.

2.1 A Conversational Unit for Fitting Data to the Weibull Distribution.

Documentation: Appendix N, Appendix Volume IV, pp. 63-101, and THEMIS Report No. 23. (Chang-Wu Yen and J. E. Norman [30])

This unit is an example of the value of a simple graphics unit to the teaching of statistics and in research. It is one of the easiest to operate (User just calls $LINK WEIBUL and answers questions as shown in Figures 2.1.1 and 2.1.2.). Three methods of estimation were employed for this rather wide class of growth curves: A modification of the quantal methods used in bio-assay, which was taken from the QUANTAL unit (Section 1.2), and two methods based on order statistics. For large samples (n = 100) it makes no difference which method is employed (see Figure 2.1.3 - the bio-assay method, and 2.1.4, all three methods).

In the experimental run of the latest version (under MVS, October 77) we employed smaller sample sizes (n = 20, bioassay method, Fig. 2.1.5) and very small sample sizes (n = 10, Fig. 2.1.6, all methods) including an example of the J-shaped members of the Weibull class (Fig. 2.1.8, n = 10, bio-assay method). The striking fact of the small sample analyses is that the numerical estimates are very different for different methods, and quite poor in relation to the parent from which sample arose (scale parameter estimate 4.33 from a population value of 2.0; shape parameter 0.41 from a parent value of .5, Figure 2.1.7) but the quality of fit is quite impressive, even for these small data (Fig. 2.1.8) and, though different growth models are fitted to the same data by different methods (Fig. 2.1.6) it is not easy to decide which one yields superior results; the main difference appears to lie in the relative importance given to the fit

```
                         OUTPUT AREA
IS PROGRAM WILL ESTIMATE THE PARAMETERS OF A TWO PARAMETER
EIBULL DISTRIBUTION USING INTERACTIVE GRAPHICAL AND MODIFIED BIOASSAY
ECHNIQUES. FOR DEMONSTRATION PURPOSES.A MONTE CARLO DATA GENERATION
OPTION IS AVAILABLE. TO USE THIS.PRESS PROGRAM KEY 5. TO ENTER REAL DATA
FROM THE KEYBOARD PRESS KEY 4
A WEIBUL CDF PLOTTER ROUTINE CAN BE REACHED BY PRESSING KEY 10.
AT ANY POINT IN THIS PROGRAM YOU MAY RESTART BY PRESSING KEY 6 OR
OR TERMINATE BY PRESSING KEY 31
```

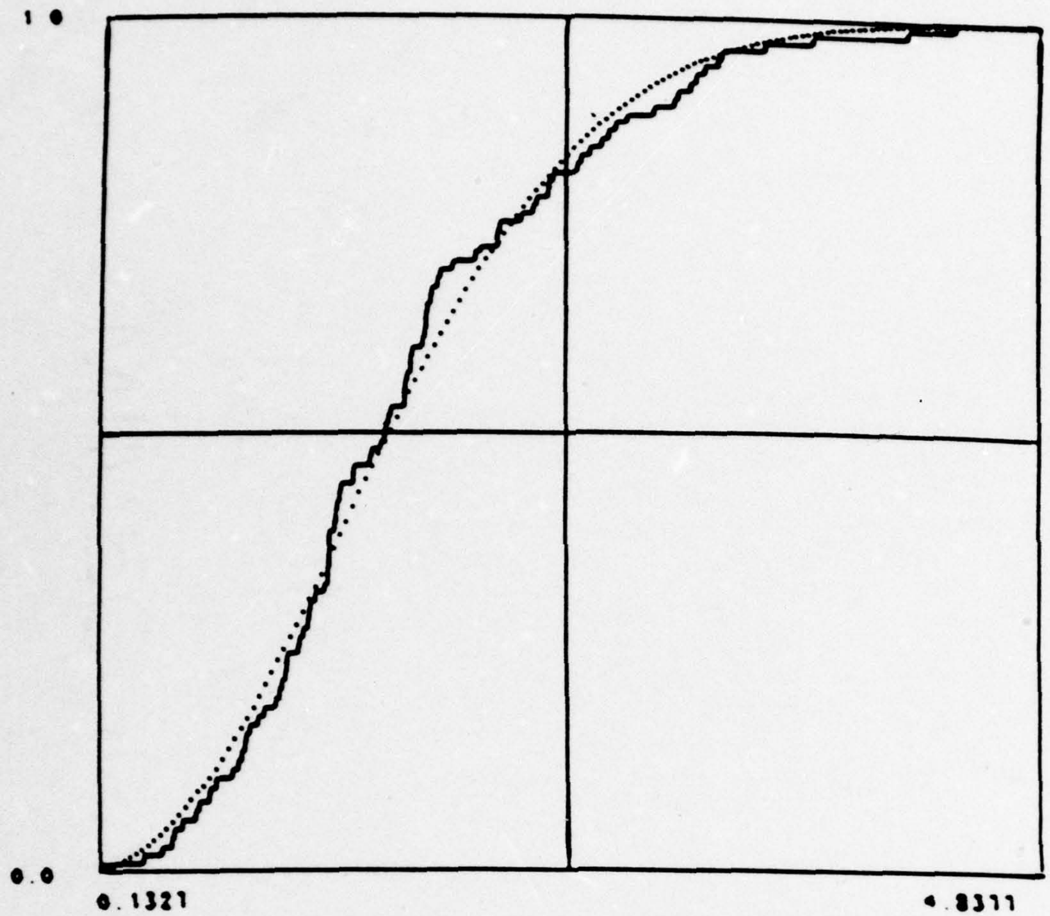
```
                         REPLY AREA
```

Fig. 2.1.1

```
                              OUTPUT AREA
 . FOLLOWING INPUT VALUES ARE REQUIRED:
 . IS THE SAMPLE SIZE.B IS THE SCALE PARAMETER.C IS THE SHAPE PARAMETER.I
 .S INDEX FOR THE THREE METHODS.
 I=1  BIOASSAY METHOD(B)
 I=2  LOG-ORDER STATISTICS(L)(BY MENON)
 I=3  ORDER STATISTICS(O)(BY BAIN AND ANTLE)
 IOD IS ANY POSITIVE ODD INTEGER USED TO GENERATE THE SAMPLE

ENTER THE VALUES OF N.I.B.C.IOD.IN THIS ORDER.SEPARTED BY COMMAS.

SAMPLE SIZE N IS 100
ACTUAL B IS    2.00000
ACTUAL C IS    2.00000
ESTIMATED B IS    1.88851
ESTIMATED C IS    2.12251

SAMPLE SKEWNESS IS      0.6715
SAMPLE KURTOSIS IS      3.4226
SKEWNESS OF ESTIMATED DENSITY IS     0.3066
KURTOSIS OF ESTIMATED DENSITY IS     3.1011
DEPRESS KEY 5 TO DISPLAY ESTIMATED WEIBULL CDF
TO OBTAIN THE CUMULATIVE PROBABILITY FOR A GIVEN VALUE OF X. DEPRESS
KEY 7.
TO OBTAIN THE PERCENTAGE POINT X FOR A GIVEN PROBABILITY. DEPRESS KEY 8.
```

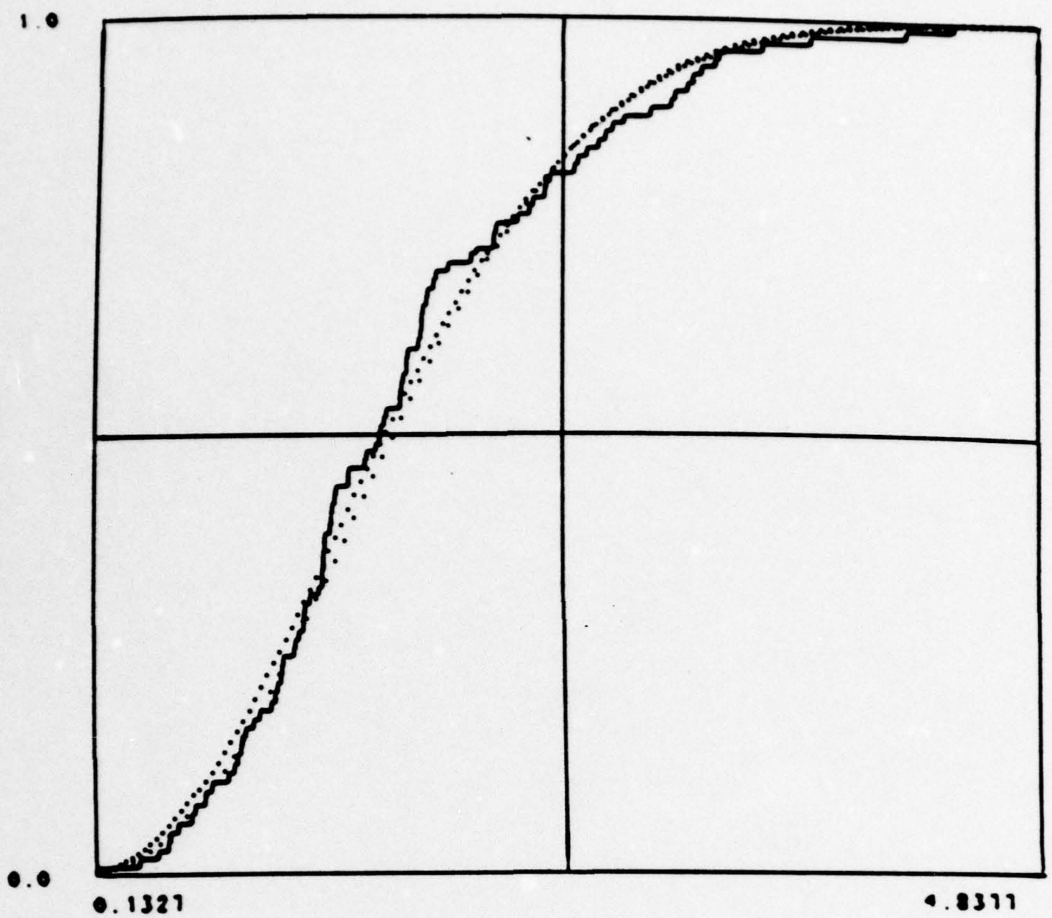
REPLY AREA

Fig. 2.1.2



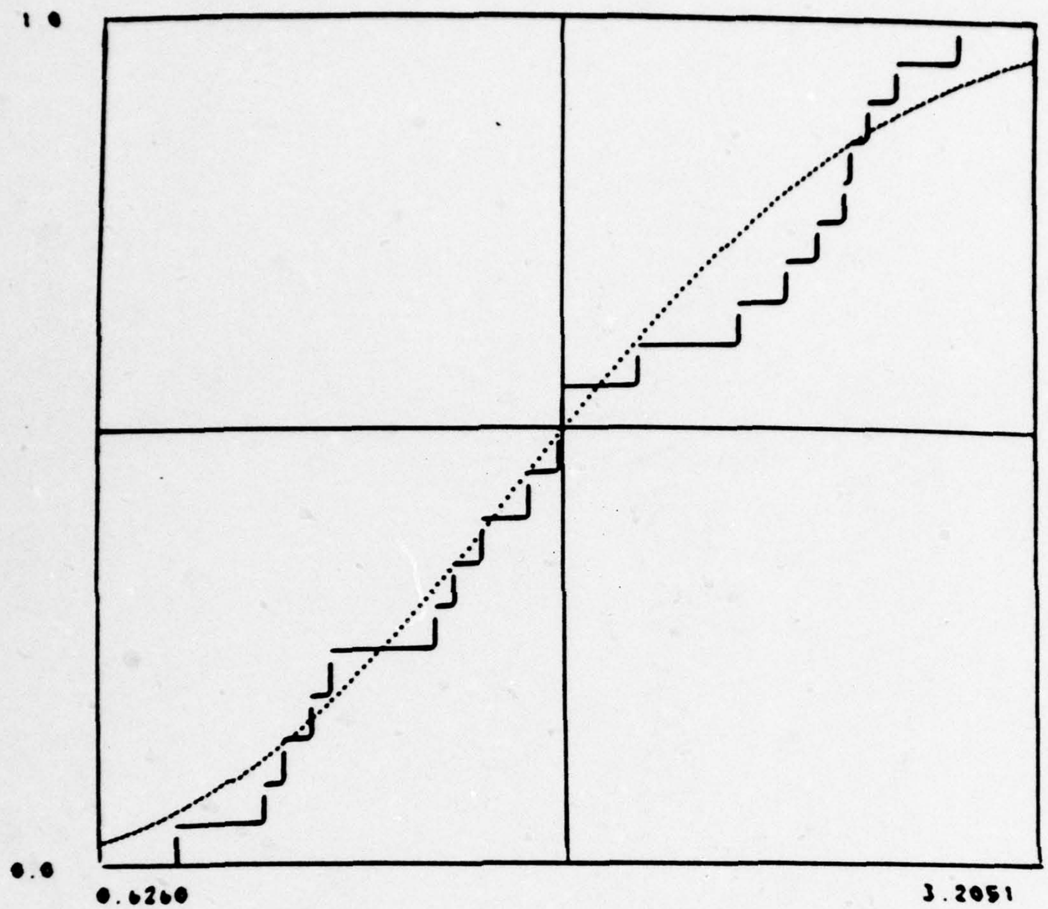
LEGEND: GRAPH OF THE ESTIMATED CDF OF FIRST METHOD USED HAS THE MOST DOTS. THAT OF THE SECOND METHOD USED HAS FEWER POINTS, AND THAT OF THE LAST METHOD HAS THE LEAST. TO USE ANOTHER METHOD PRESS KEY 10

Fig. 2.1.3



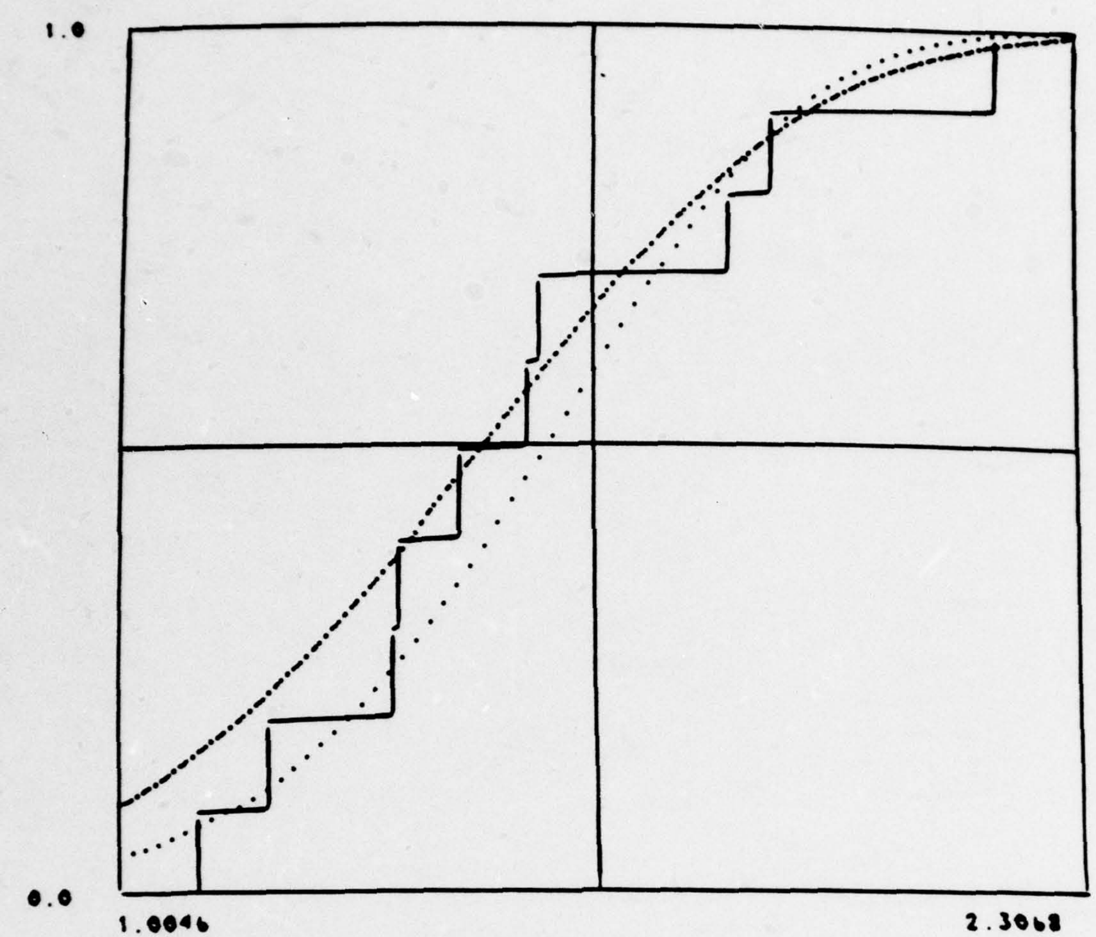
LEGEND: GRAPH OF THE ESTIMATED CDF OF FIRST METHOD USED HAS THE MOST DOTS. THAT OF THE SECOND METHOD USED HAS FEWER POINTS. AND THAT OF THE LAST METHOD HAS THE LEAST. TO USE ANOTHER METHOD PRESS KEY 10

Fig. 2.1.4



LEGEND: GRAPH OF THE ESTIMATED CDF OF FIRST METHOD USED HAS THE MOST DOTS. THAT OF THE SECOND METHOD USED HAS FEWER POINTS. AND THAT OF THE LAST METHOD HAS THE LEAST. TO USE ANOTHER METHOD PRESS KEY 10

Fig. 2.1.5



LEGEND: GRAPH OF THE ESTIMATED CDF OF FIRST METHOD USED HAS THE MOST DOTS. THAT OF THE SECOND METHOD USED HAS FEWER POINTS. AND THAT OF THE LAST METHOD HAS THE LEAST. TO USE ANOTHER METHOD PRESS KEY 10 TO SEE A MAGNIFIED PORTION OF THIS GRAPH,PRESS KEY 11

Fig. 2.1.6

```
                         OUTPUT AREA
  FOLLOWING INPUT VALUES ARE REQUIRED:
. IS THE SAMPLE SIZE.B IS THE SCALE PARAMETER.C IS THE SHAPE PARAMETER.I
IS INDEX FOR THE THREE METHODS.
I=1  BIOASSAY METHOD(B)
I=2  LOG-ORDER STATISTICS(L)(BY MENON)
I=3  ORDER STATISTICS(O)(BY BAIN AND ANTLE)
IOO IS ANY POSITIVE ODD INTEGER USED TO GENERATE THE SAMPLE

ENTER THE VALUES OF N.I.B.C.IOO.IN THIS ORDER.SEPARTED BY COMMAS.

SAMPLE SIZE N IS  10
ACTUAL B IS    2.00000
ACTUAL C IS    0.50000
ESTIMATED B IS    4.32970
ESTIMATED C IS    0.40101

SAMPLE SKEWNESS IS     0.7140
SAMPLE KURTOSIS IS     2.5163
SKEWNESS OF ESTIMATED DENSITY IS  117.4499
KURTOSIS OF ESTIMATED DENSITY IS  261.6807
DEPRESS KEY 5 TO DISPLAY ESTIMATED WEIBULL CDF
TO OBTAIN THE CUMULATIVE PROBABILITY FOR A GIVEN VALUE OF X, DEPRESS
KEY 7.
TO OBTAIN THE PERCENTAGE POINT X FOR A GIVEN PROBABILITY, DEPRESS KEY 8.

                         REPLY AREA
```

Fig. 2.1.7

LEGEND: GRAPH OF THE ESTIMATED CDF OF FIRST METHOD USED HAS THE MOST DOTS. THAT OF THE SECOND METHOD USED HAS FEWER POINTS. AND THAT OF THE LAST METHOD HAS THE LEAST. TO USE ANOTHER METHOD PRESS KEY 10.

Fig. 2.1.8

of the growth model in the "tails," a well known characteristic in the analysis of growth or learning curves.

Novel features in this unit are:

The use of a very efficient modified bio-assay method for fitting a complex class of growth models, and the comparison with the more generally used order statistics methods. A facility to obtain fits to data by different methods and making a visual comparison, rather than relying on numerical values of scale and shape parameters; frequently, widely differing parameter estimates yield equally good fits.

This unit has been used in classroom instruction (courses in Distribution Theory, and Non-linear Statistical Analysis) and for illustration and verification of dissertation work on the topic of estimation of parameters of the Weibull class. The unit has been shown to be operational and transportable (under MVS) IBM 370/158.

2.2 Conversational Unit for Spline Function Construction.

Documentation: Appendix O, Appendix vol. IV, pp. 102-137, and Themis Report No. 24 (J. S. Scott and J. E. Norman [31]).

Data filtering and smoothing is an important preparatory step in the analysis of time series or fitting of densities fo mixed statistical distributions. In the first step, some authors begin the process* by fitting a cubic spline to the observed data, and then apply various techniques of filtering (or moving averages) to obtain a smoother (monotonic, unimodal, bimodal) fit.

A graphical unit seems especially useful to facilitate this process. In the present unit, after calling $LINK SPLINE, the user receives instruction to define coordinates and number of points, and then to enter the points (Fig. 2.2.1 and 2.2.2). The program fits a cubic spline through these points (Fig. 2.2.3), then enables the user to change ordinates of any point or points by the use of a light pen. In Fig 2.2.4 the light pen was placed on the asterisk near the center of the field, and a series of dots was erected around this point.

---

*e.g., G. Wahba, "Smoothing," Symp. Appl. Stat., Dayton, 1976 (proceedings in print).

OUTPUT AREA

INSTRUCTIONS - PLEASE READ CAREFULLY

YOU ARE NOW ENTERING THE SPLINE PROGRAM.

YOU WILL NOW BE ASKED SEVERAL QUESTIONS. YOU WILL INDICATE YOUR ANSWERS BY TYPING THEM ON THE TYPEWRITER KEYBOARD IN FRONT OF YOU

WHEN YOU HAVE COMPLETED YOUR ANSWER, YOU MUST INDICATE THIS BY PRESSING THE "ALT" KEY, AND WHILE HOLDING IT DOWN, PRESS THE "S" KEY AFTER YOU HAVE SIGNALLED YOUR COMPLETION OF AN ANSWER, YOUR REPLY WILL BE DISPLAYED BACK TO YOU ON THE SCREEN. AT THIS TIME YOU WILL HAVE THE OPPORTUNITY TO CHANGE OR RE-ENTER YOUR ANSWER. IF YOU DESIRE TO CHANGE YOUR REPLY, PRESS PROGRAM FUNCTION KEY 2; OTHERWISE PRESS PROGRAM FUNCTION KEY 1, AND THE PROGRAM WILL CONTINUE.

AT ANY POINT IN THIS PROGRAM YOU MAY RESTART BY PRESSING PROGRAM FUNCTION KEY 30 OR TERMINATE BY PRESSING PROGRAM FUNCTION KEY 31

PLEASE BE SURE THAT YOU ARE FAMILIAR WITH ALL OF THE ABOVE INSTRUCTIONS. PRESS KEY 1 WHEN YOU ARE READY TO CONTINUE.

Fig. 2.2.1

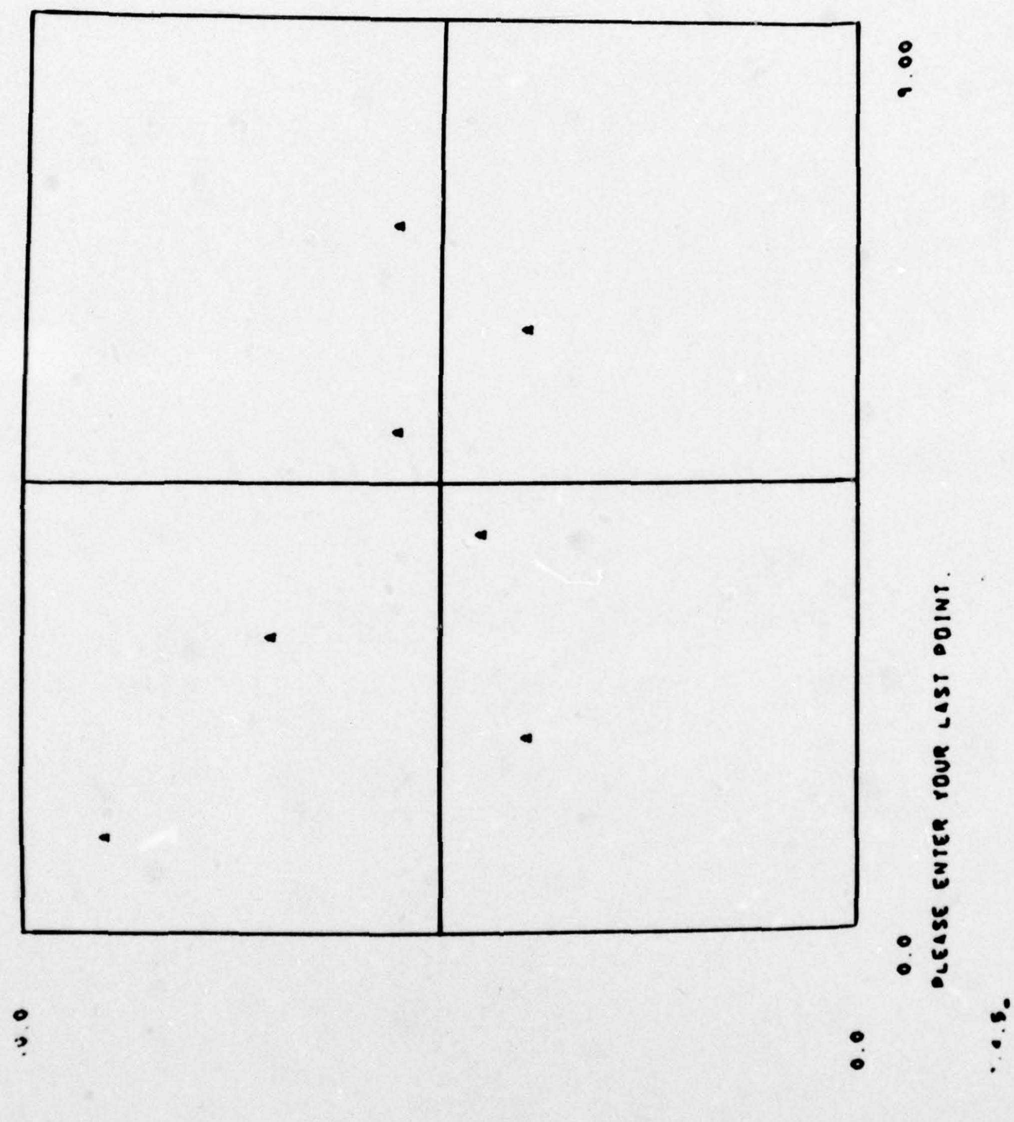


Fig. 2.2.2

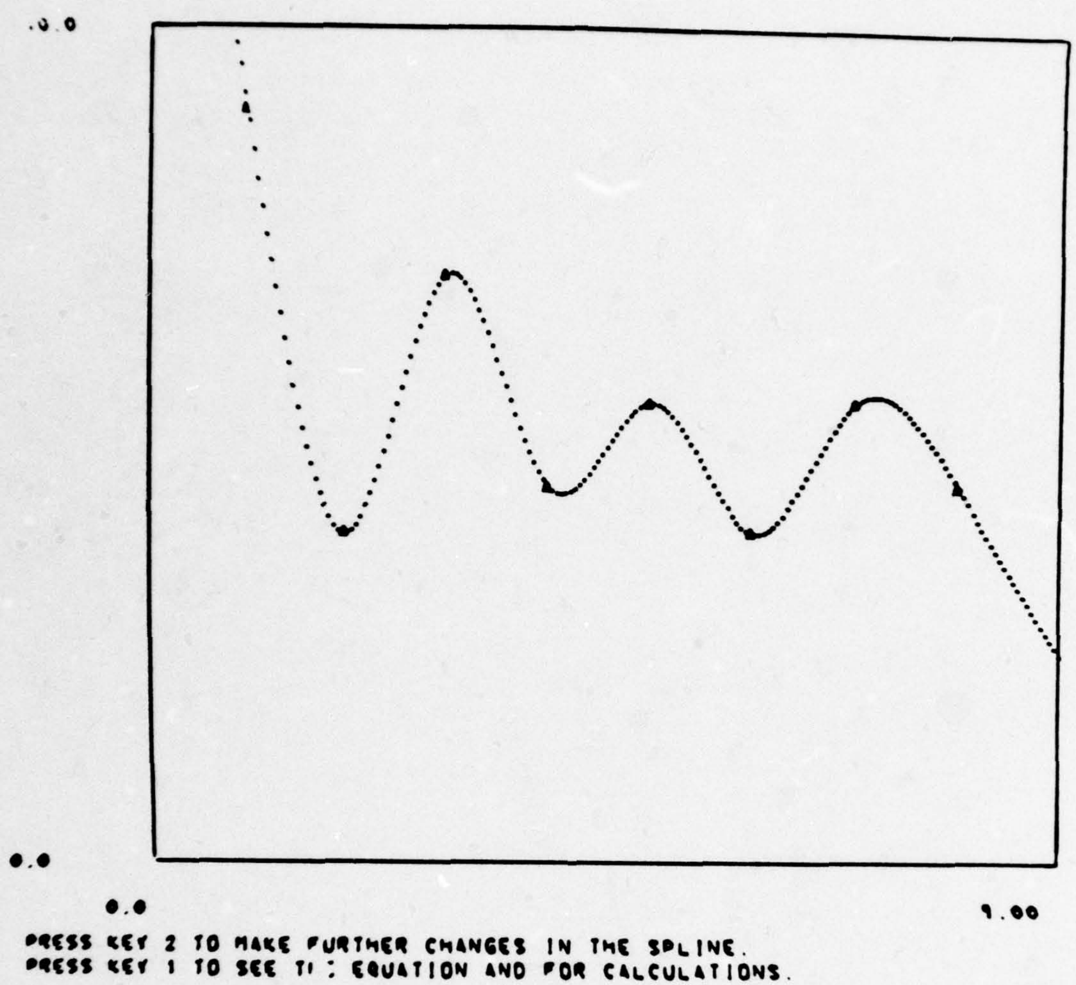

0.0
0.0
1.00
PRESS KEY 2 TO MAKE FURTHER CHANGES IN THE SPLINE.
PRESS KEY 1 TO SEE TI : EQUATION AND FOR CALCULATIONS.

Fig. 2.2.3

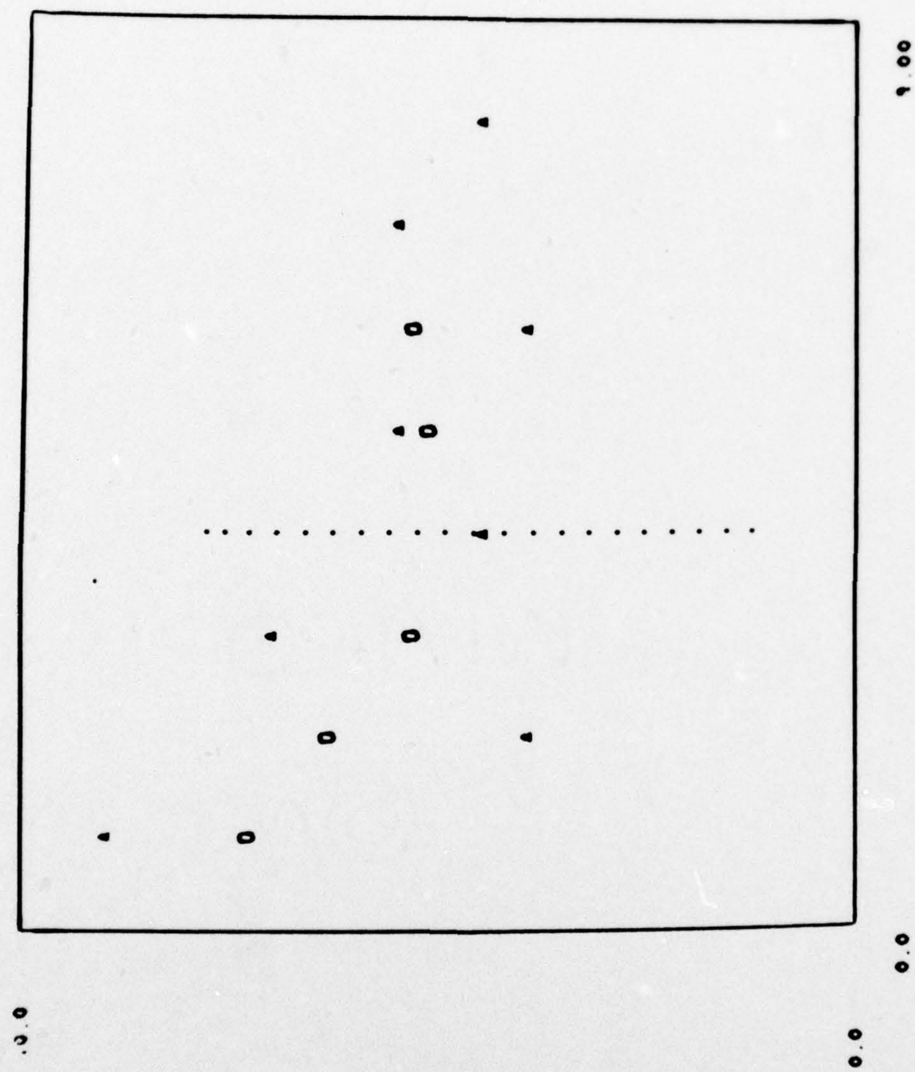


Fig. 2.2.4

Pointing the light pen at any of the dots, the user can produce an 0 at this point indicating the y-value through which the smoothed spline would be fitted. Fig. 2.2.5 shows how the smoothed spline is passed through the 0's and unchanged asterisks. Another illustration is in Fig. 2.2.6 and 2.2.7.

After the decision has been made (after as many changes as desired), the user obtains the equation of the last spline (in the usual truncated polynomial notation; the small illegible mark before the **3 in each row of Fig. 2.2.8 is a +), and may use this for purposes of interpolation or estimation of trends, or parameters of mixed distributions.

Novel features: None really; merely a fast routine to draw splines throught points relocated by the use of a light pen.

The illustration 2.2.1 to 2.2.8 were photographs of the screen, obtained in October, 1977, with the IBM 2250/2840 hooked up to a 370/158 under MVS. It also operated well after transportation of all data sets.

Use has been made of this unit in some courses, and for the smoothing of data prior to performing a compartmental analysis (estimation of coefficients in a system of differential equations) in a dissertation study of tracers.

2.3. An Interactive Analysis-of-Covariance Unit

Documentation: Appendix P, Appendix Vol. IV, pp. 138-225, and THEMIS Report No. 25 (M. E. Nash and R. E. Bargmann [33]).

Programs performing Analysis of Variance, of a two-way classification model, even with missing and unbalanced cells, have been in more than adequate supply for two decades. On the other hand, analysis-of-covariance units, except for the very simplest cases, are in shorter supply. One reason is the problem of what to regard as classificatory variables (the number of rows and columns in a design) and what to regard as concomitant variables (design variables known without errror, but with ordinal scale characteristics). Most packages provide a "general linear analysis" feature, to deal with the more complex

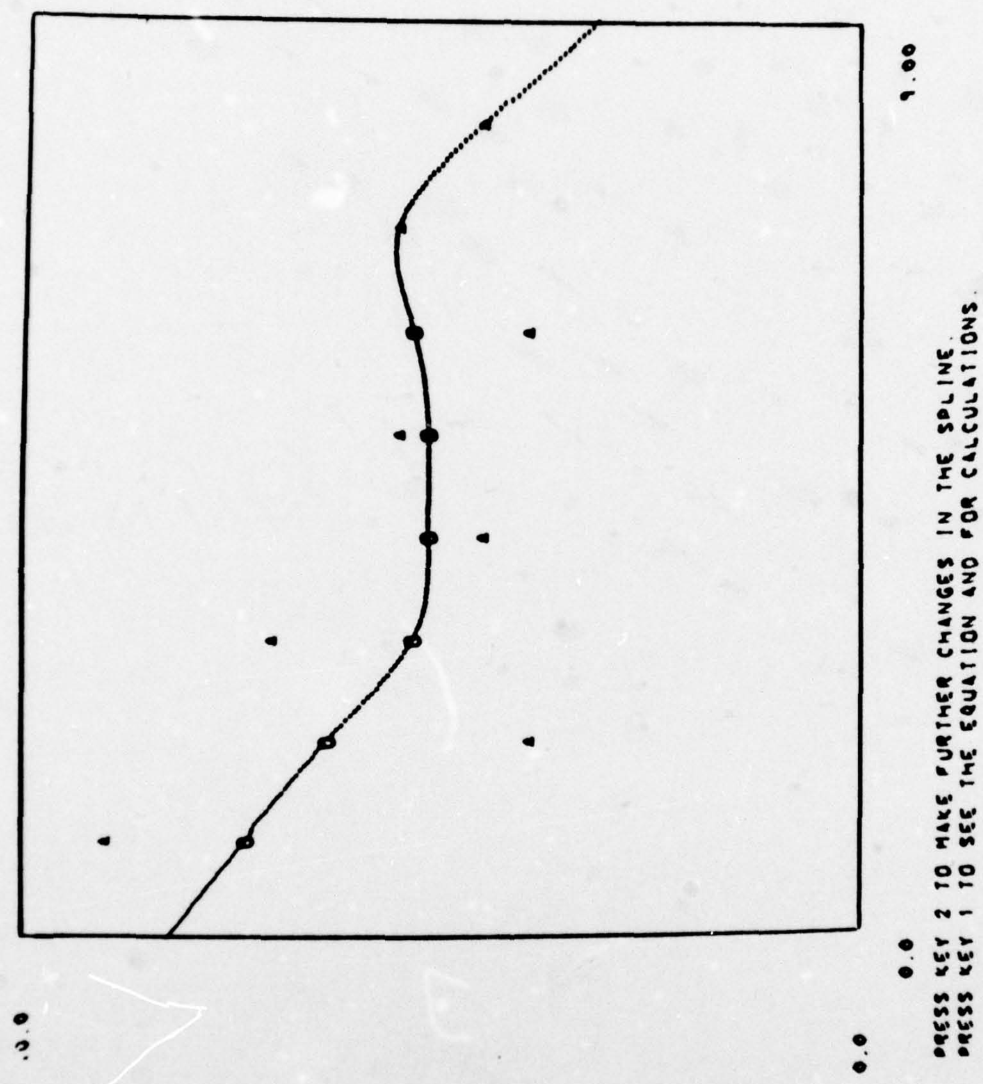

0.0
1.00
0.0
PRESS KEY 2 TO MAKE FURTHER CHANGES IN THE SPLINE.
PRESS KEY 1 TO SEE THE EQUATION AND FOR CALCULATIONS.

Fig. 2.2.5

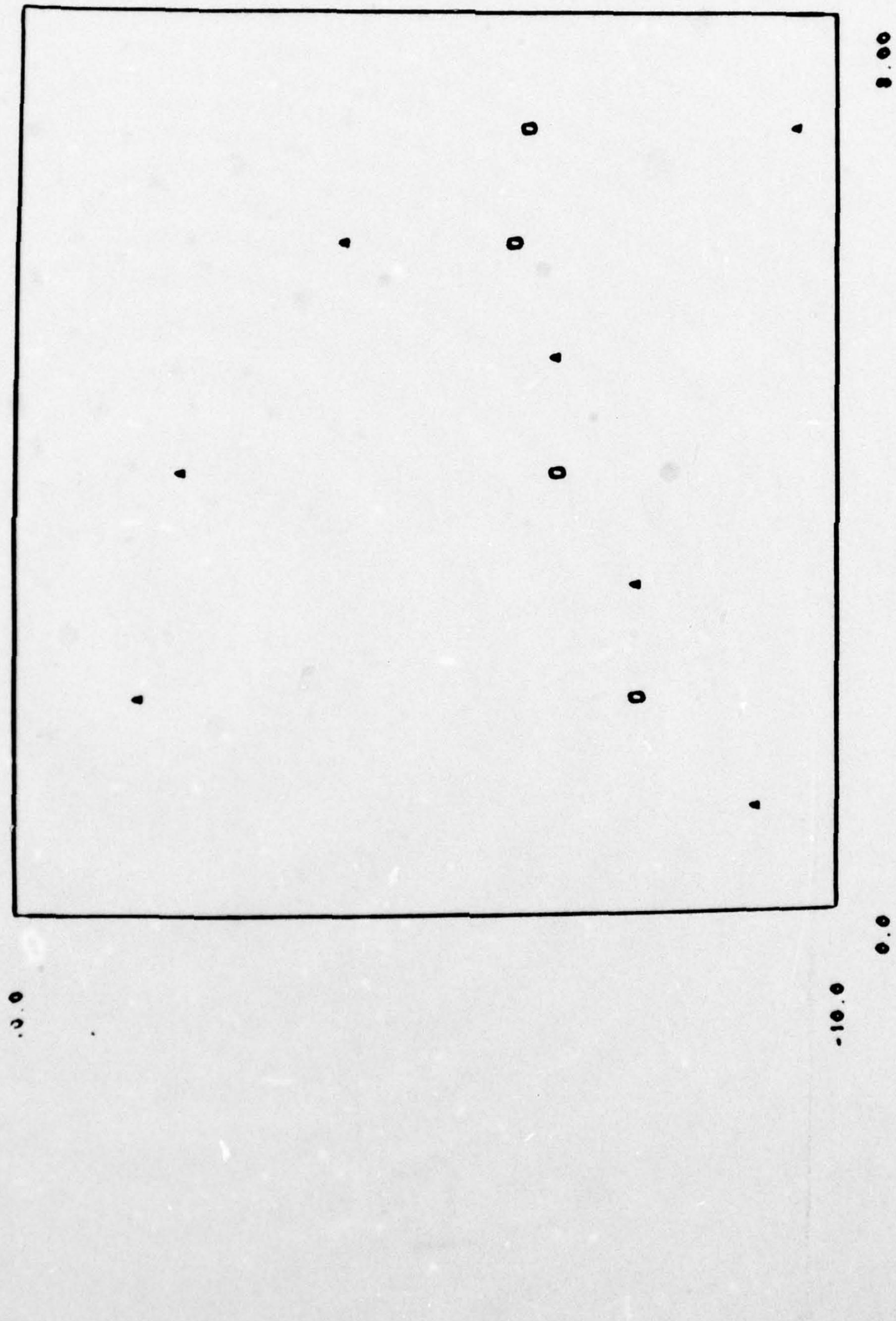


Fig. 2.2.6

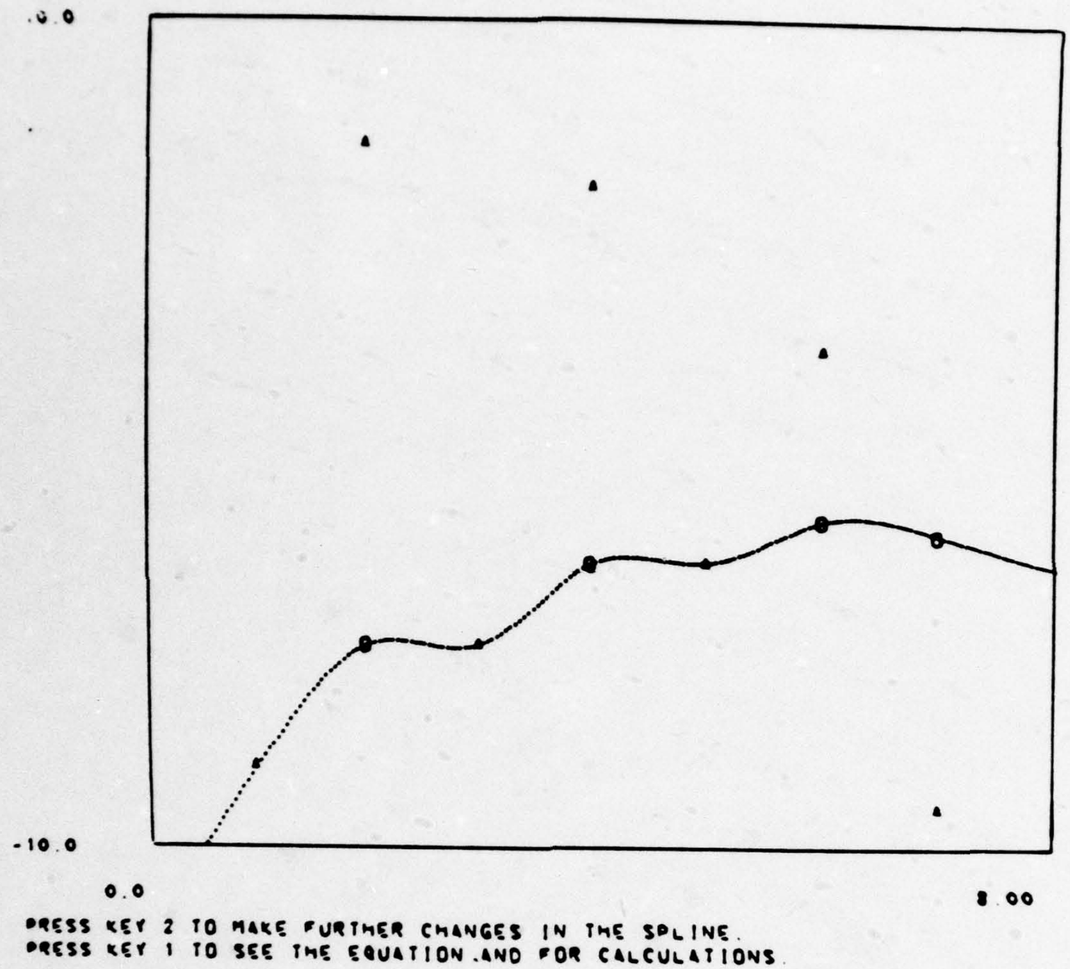

-10.0
0.0
8.00
PRESS KEY 2 TO MAKE FURTHER CHANGES IN THE SPLINE.
PRESS KEY 1 TO SEE THE EQUATION.AND FOR CALCULATIONS.

Fig. 2.2.7

OUTPUT AREA

```
Y =   8.261508180 00 +  -9.404218810-01X
    +  -5.605511310-02(X -  1.000000000 00)+ AA3
    +   3.363306820-01(X -  2.000000000 00)+ AA3
    +  -5.246896100-01(X -  3.000000000 00)+ AA3
    +   2.910118200-01(X -  4.000000000 00)+ AA3
    +  -1.851448990-02(X -  5.000000000 00)+ AA3
    +  -4.085286320-01(X -  6.000000000 00)+ AA3
    +   6.825835060-01(X -  7.000000000 00)+ AA3
    +  -3.080181030-01(X -  8.000000000 00)+ AA3
PRESS KEY 1 IF YOU WISH TO USE THIS EQUATION FOR CALCULATIONS.
```

REPLY AREA

Fig. 2.2.8

situations of analysis of covariance, but fail to warn the user that, where variables are categorized or "nominal", widely different results can be obtained by a change of assumptions on the nature of main effects and interactions, in the unbalanced case.

An interactive unit to perform an analysis of covariance involving unbalanced designs, with graphical display and variable selection features, seemed to serve a useful purpose. This unit was implemented in 1973, on the Graphics terminal system (later adapted to teletype use, see Section 3.3). When the user calls $LINK ANACOV, he receives rather extensive instructions on the screen, first to set up his data for an analysis of variance (without regression variables present). The user determines the number of rows, and columns, indicating how he wishes to encode the levels of his factors, and then proceeds to enter the data (Fig. 2.3.2). There are facilities for editing and augmenting data. Up to 9 concomitant variables can be added; if there are fewer, the last variable can be expanded into a polynomial of higher degree, or the last two can be expanded to include mixed polynomials. Figure 2.3.2 is an example of just two concomitant variables (named RAIN and SUN) with the program filling in the values for all quadratic and cubic terms in these variables. Transformations on the random variables, and on the concomitant variables, can then be indicated (Fig. 2.3.3).

The program returns with a standard analysis of variance: Tables of cell means, adjusted means for the levels of each factor (here called LOCA and FERT) and the usual analysis-of-variance table of mean square and F ratios (Fig. 2.3.4 and 2.3.5). At this stage the user is encouraged to edit his data, he may even plot the data from selected cells, rows, or columns (the random variable vs. each selected concomitant variable) in order to detect outliers.

Up to this stage, data could have been entered in batch mode - in fact, there is the option to do so. The purely conversational part begins here with the names of the concomitant variables appearing, one by one, and the user deciding, by pressing Program Function Key 1 or 2, whether he wishes to retain or exclude that variable in this pass. Figure 2.3.6 indicates that here the variable RAIN, SUN, and the square and cube of the first were to be retained. In a later pass, only RAIN, its square and cube were selected.

OUTPUT AREA

THIS PROGRAM IS DESIGNED TO PERFORM AN ANALYSIS OF COVARIANCE. YOU HAVE ONE RESPONSE VARIABLE, TWO FACTORS, AND A MAXIMUM OF NINE CONCOMITANT VARIABLES. IF YOU HAVE 8 OR FEWER CONCOMITANT VARIABLES, YOU MAY HAVE A POLYNOMIAL GENERATED IN THE LAST ONE YOU LIST. IF YOU HAVE 2 CONCOMITANT VARIABLES YOU MAY HAVE QUADRATIC AND CUBIC TERMS INVOLVING THEM GENERATED FOR YOU. IF YOU HAVE 3 CONCOMITANT VARIABLES YOU MAY HAVE QUADRATIC TERMS INVOLVING THOSE 3 VARIABLES GENERATED.

YOU WILL BE GIVEN OPPORTUNITIES TO SEE PLOTS OF YOUR DATA OR TO SEE THE RAW DATA AGAIN. THUS, YOU WILL BE ABLE TO EDIT YOUR DATA AND RUN THE ANALYSES AGAIN.

TO BEGIN YOU MUST ANSWER QUESTIONS BY USING THE TYPEWRITER KEYBOARD DIRECTLY IN FRONT OF YOU. TO SIGNAL YOUR COMPLETION OF QUESTIONS, FIRST DEPRESS THE "ALT" KEY, AND WHILE HOLDING IT DOWN, DEPRESS THE "5" KEY. THIS SEQUENCE WILL LATER BE REFERRED TO AS "EOB". ONCE YOU ANSWER QUESTIONS, THE ANSWERS WILL BE DISPLAYED BACK TO YOU. IF YOU ARE NOT SATISFIED, PRESS KEY 2 TO REENTER DATA. OTHERWISE, THE PROGRAM WILL CONTINUE BY YOUR PRESSING ANY KEY.

AT ANY TIME YOU MAY RESTART BY PRESSING KEY 30 OR TERMINATE BY PRESSING KEY 31.

CAUTION: DO NOT TRY TO SPEED UP THE PROGRAM BY ANSWERING QUESTIONS BEFORE THEY ARE ASKED. THIS WILL ONLY CREATE PROBLEMS.

IF YOU HAVE PREVIOUSLY USED THIS PROGRAM OR ENTERED YOUR DATA THROUGH BATCH MODE, PRESS KEY 2 TO SEE YOUR DATA.

PRESS KEY 1 TO PROCEED.

REPLY AREA

Fig. 2.3.1

OUTPUT AREA

	LOC	FERT	YILD KKAY	RAIN KAYY	SUN KAY	KAA2 YAA2	KAA3 YAA3
1	1	1	50.0000	10.0000	5.00000	100.000	1000.00
			500.000	250.000	50.0000	25.0000	125.000
2	1	1	60.0000	9.00000	7.00000	81.0000	729.000
			567.000	441.000	63.0000	49.0000	343.000
3	1	1	60.0000	9.00000	6.00000	81.0000	729.000
			486.000	324.000	54.0000	36.0000	216.000
4	1	2	40.0000	10.0000	3.00000	100.000	1000.00
			300.000	90.0000	30.0000	9.00000	27.0000
5	1	2	30.0000	11.0000	4.00000	121.000	1331.00
			484.000	176.000	44.0000	16.0000	64.0000
6	1	3	80.0000	5.00000	9.00000	25.0000	125.000
			225.000	405.000	45.0000	81.0000	729.000
7	1	3	82.0000	4.00000	9.00000	16.0000	64.0000
			144.000	324.000	36.0000	81.0000	729.000
8	2	1	40.0000	10.0000	5.00000	100.000	1000.00
			500.000	250.000	50.0000	25.0000	125.000
9	2	1	50.0000	11.0000	6.00000	121.000	1331.00
			126.000	396.000	66.0000	36.0000	216.000
10	2	1	55.0000	12.0000	6.00000	144.000	1728.00
			864.000	432.000	72.0000	36.0000	216.000
11	2	2	35.0000	12.0000	2.00000	144.000	1728.00
			288.000	48.0000	24.0000	4.00000	8.00000
12	2	2	37.0000	11.0000	4.00000	121.000	1331.00
			484.000	176.000	44.0000	16.0000	64.0000
13	2	3	80.0000	3.00000	7.00000	9.00000	27.0000
			63.0000	147.000	21.0000	49.0000	343.000
14	2	3	83.0000	4.00000	12.0000	16.0000	64.0000
			192.000	576.000	48.0000	144.000	1728.00
15	2	3	83.0000	4.00000	11.0000	16.0000	64.0000
			176.000	484.000	44.0000	121.000	1331.00
16	2	3	85.0000	6.00000	10.0000	36.0000	216.000
			360.000	600.000	60.0000	100.000	1000.00

PRESS KEY 1 TO CONTINUE OR KEY 2 TO CHANGE DATA.

IF YOU WISH TO SEE PLOTS OF YOUR DATA, PRESS KEY 6.

IF, AT ANY TIME, YOU WISH TO RETURN TO YOUR PREVIOUS ANALYSIS OUTPUT, PRESS KEY 29.

REPLY AREA

Fig. 2.3.2

```
                              OUTPUT AREA
CODE      TRANSFORMATION
0.1       NO TRANSFORMATION
  2       LOGE(X)    (X>0)
  3       LOGE(1+X)    (X>-1)
  4       SQRT(X)    (X>=0)
  5       1/X    (X>0)
  6       ARCSIN(2X-1)    (0<X<1)
          VARIANCE-STABILIZING TRANSFORMATION FOR PROPORTIONS
  7       AVAILABLE (NOW: NO TRANSFORMATION)

SELECT FROM THE LIST ABOVE A TRANSFORMATION CODE FOR EACH OF THE
VARIABLES.  TYPE THE CODE FOR THE RESPONSE VARIABLE, FOLLOWED BY
A CODE FOR EACH CONCOMITANT VARIABLE.  SEPARATE ALL ANSWERS BY COMMAS.

 YILD  RAIN  SUN
  1     1     1
```

```
                              REPLY AREA
```

Fig. 2.3.3

```
                              OUTPUT AREA
                                                                PAGE   2L
YIELD AS A FUNCTION OF RAIN AND SUN
                           VARIABLE = YILD
              TABLE OF MEANS AND EFFECTS    LOCA VERSUS FERT

                   ROWS = LOCA        COLUMNS = FERT
                 ROWS 1, 2, AND 3 DENOTE CELL MEAN, SIZE,
                 AND STANDARD DEVIATION RESPECTIVELY.
CODED
LEVEL     1            2            3

     5.667E 01    3.500E 01    8.100E 01    5.743E 01
 1       3            2            2            7
     5.774E 00    7.071E 00    1.414E 00

     4.833E 01    3.600E 01    8.275E 01    6.089E 01
 2       3            2            4            9
     7.638E 00    1.414E 00    2.062E 00

COL  5.250E 01    3.550E 01    8.217E 01    5.938E 01
         6            4            6           16

                ESTIMATES OF LOCA ADJUSTED MEANS, ORDERED

         1            2
     6.072E 01    5.833E 01

                ESTIMATES OF FERT ADJUSTED MEANS, ORDERED

         3            1            2
     8.242E 01    5.235E 01    3.535E 01

                              REPLY AREA
```

Fig. 2.3.4

OUTPUT AREA

PAGE 3R

ANALYSIS OF VARIANCE FOR VARIABLE YILD

SOURCE OF VARIATION	MEAN SQUARE	F VS. ERROR	F VS. INTERACTION
LOCA	2.192026E 01	8.765E-01	5.020E-01
FERT	2.827594E 03	1.131E 02	6.476E 01
INTERACTION	4.366487E 01	1.746E 00	
SUBTOTAL	1.157933E 03	4.630E 01	
ERROR	2.500833E 01		
TOTAL	4.026499E 02		

ROOT MEAN SQUARE ERROR = 5.00083

ROOT MEAN SQUARE INTERACTION = 6.60794

REPLY AREA

Fig. 2.3.5

OUTPUT AREA

YOUR CONCOMITANT VARIABLE NAMES WILL APPEAR BELOW ONE AT A TIME.
PRESS KEY 1 IF YOU WISH TO INCLUDE THE VARIABLE OR KEY 2 IF
YOU WISH TO DELETE THE VARIABLE.

RAIN
SUN
KAA2
KAA3

THE ABOVE VARIABLES ARE THE ONES YOU DECIDED TO INCLUDE.
IF YOU AGREE, PRESS KEY 1: IF YOU WOULD LIKE TO TRY AGAIN,
PRESS KEY 2.

REPLY AREA

Fig. 2.3.6

The program returns the matrices of sums of squares and products for error, the correlation based on them, and the regression coefficients, for the selected variables, (Fig. 2.3.7). Plots of all data (Fig. 2.3.8) or selected rows or columns (Fig. 2.3.9) can be viewed, with the regression polynomial displayed. Notice that the polynomial applies to the data "within cells", and total data, or total rows, would be affected by main effects in addition to regression on concomitant variables. An analysis-of-covariance table (Fig. 2.3.10) can be obtained, at any time, and the user can decide if another pass with another selection of concomitant variables and/or data editing are desired.

Novel features: Analysis of covariance with unbalanced designs, with plotting facilities, variable selection, and expansion of variables into polynomials.

This unit, and its later adaptation to teletype terminals (attached to CDC Cyber) has been in frequent use in the classroom. It was demonstrated to visiting groups, reported on in meetings of the American Statistical Association, and demonstrated at locations away from the University of Georgia, over remote job entry or telephone lines (Armstrong College and West Georgia College, 1975). It was shown to be operational and transportable under MVS, in October and November, 1977; the attached figures are photographs of the screen from these trials.

```
                              OUTPUT AREA
                                                          PAGE    4L
YIELD AS A FUNCTION OF RAIN AND SUN
                           ANALYSIS OF COVARIANCE
                MATRIX E AFTER ELIMINATION OF FACTORS(LOCA) AND (FERT)
              YILD          RAIN          KAA2          KAA3
YILD   250.083       8.58333       128.917       1929.08
              YILD          RAIN          KAA2          KAA3
RAIN   8.58333       8.91667       126.917       1615.42
              YILD          RAIN          KAA2          KAA3
KAA2   128.917       126.917       2141.58       30711.4
              YILD          RAIN          KAA2          KAA3
KAA3   1929.08       1615.42       30711.4       471291.
                         CORRELATIONS BASED ON E
              YILD          RAIN          KAA2          KAA3
YILD   1.00000       0.181766      0.176157      0.177690
              YILD          RAIN          KAA2          KAA3
RAIN 0.181766        1.00000       0.918438      0.788023
              YILD          RAIN          KAA2          KAA3
KAA2 0.176157        0.918438      1.00000       0.966691
              YILD          RAIN          KAA2          KAA3
KAA3 0.177690        0.788023      0.966691      1.00000
                         REGRESSION COEFFICIENTS
              RAIN          KAA2          KAA3
       37.9908       -5.83967      0.254413

                               REPLY AREA
```

Fig. 2.3.7

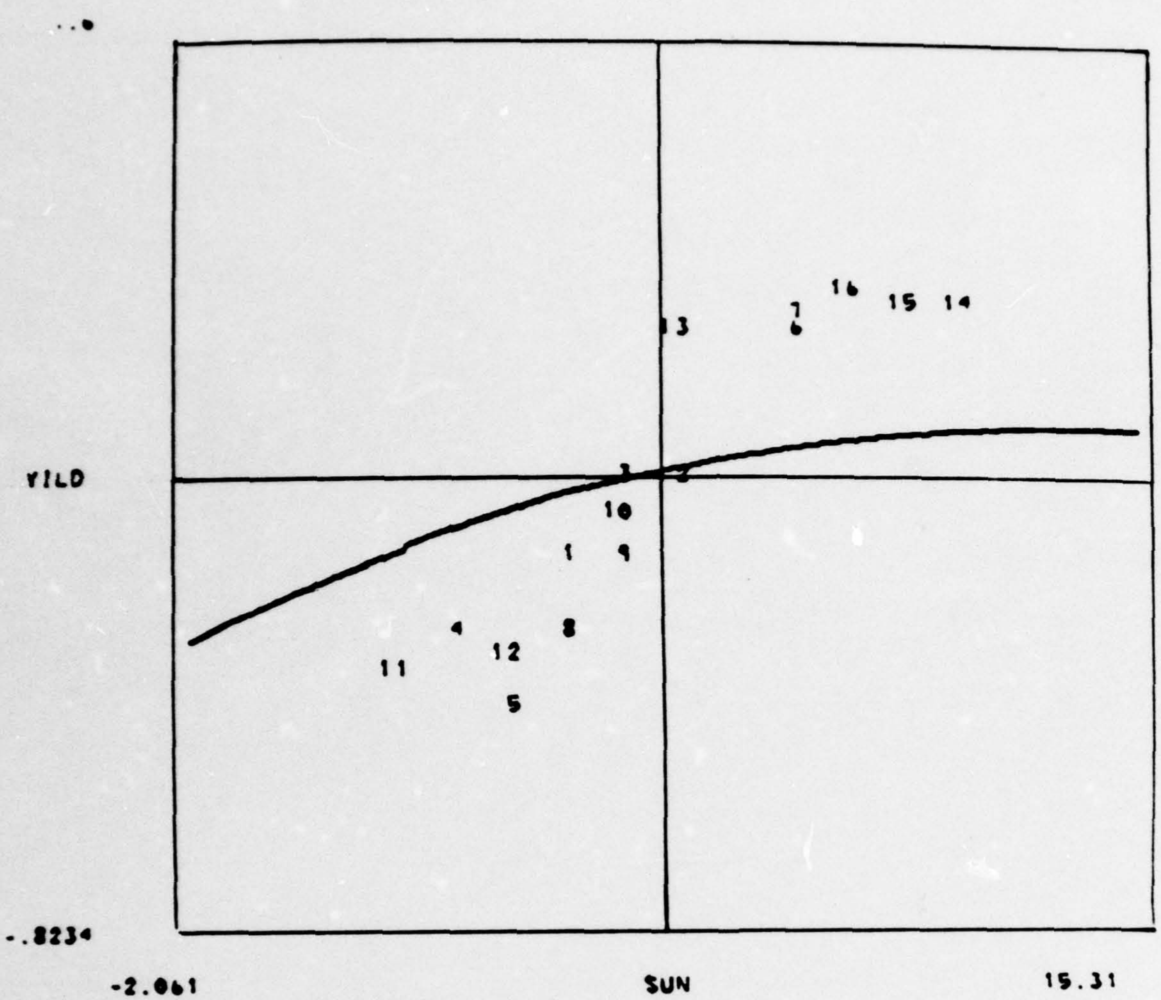


Fig. 2.3.8

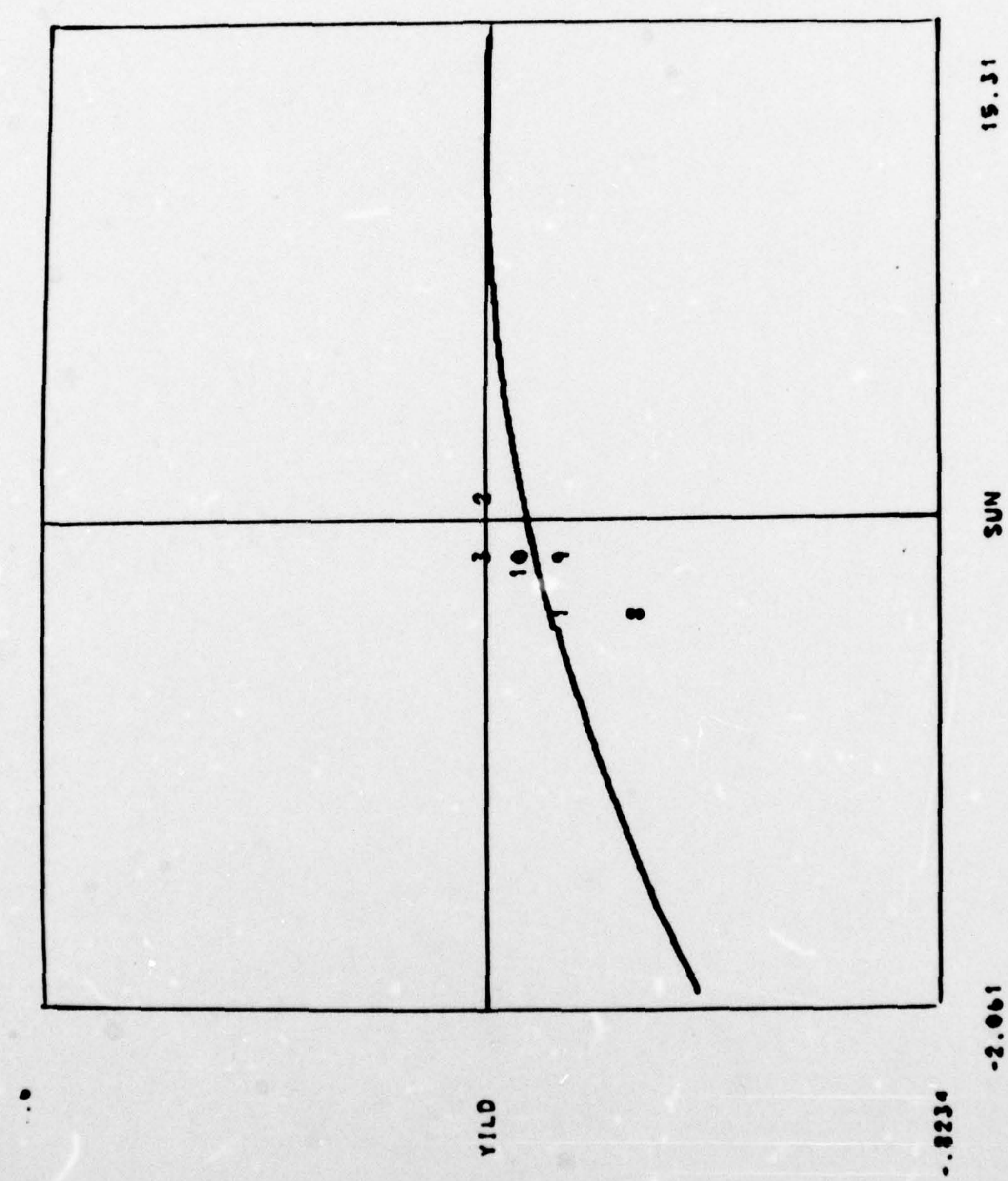


Fig. 2.3.9

OUTPUT AREA

PAGE 6R

ANALYSIS OF COVARIANCE

SOURCE OF VARIATION	MEAN SQUARE	F VS. ERROR	F VS. INTERACTION
LOCA	4.005447E 00	1.507E-01	1.274E-01
FERT	2.183695E 02	8.216E 00	6.943E 00
INTERACTION	3.145203E 01	1.183E 00	
SUBTOTAL	9.639014E 01	3.627E 00	
ERROR	2.657150E 01		

ROOT MEAN SQUARE ERROR = 5.15534

MEAN SQUARE REGRESSION = 21.3469

REPLY AREA

Fig. 2.3.10

2.4. <u>An Interactive Multivariate Data Analysis Program.</u>

Documentation: Appendix J, Appendix vol. III, pp. 72-209, and THEMIS Report No. 18 (A. Ballengee and R. E. Bargmann [24])

The structure of this unit is so similar to the preceding (Analysis of Covariance) unit that a separate set of photographs was not included here; a very detailed sample of displays is to be found in Appendix J. The unit was fully operational in our October/November 1977 trials.

The characterististics of experimental design are the same in this unit and the preceding one: A one-way classification or two-way classification design, with possibly missing cells and unbalanced data. The user may encode the levels of his factors, and may specify transformation of data. Runs can be repeated as often as desired with changed transformations. There is a facility for data-editing, and data may be entered through batch mode.

After data entry is complete, the program performs an analysis of variance on each response variable separately. Since there is no distinction between variables into random and concomitant variables here, the user may stop the univariate analysis at any time. Usually, however, the user would let the analysis of variance be completed before proceeding to variable selection for the multivariate pass. He may inspect tables of cell means and standard deviations within cells, adjusted means for rows and columns (in the case of imbalance) and the usual analysis of variance tables.

The multivariate pass begins with variable selection - the user may decide to include any variable whose name appears on the screen (by depressing Program Function Key 1) or ignore it (by depressing Program Function Key 2). Such decision is usually based on the F-values observed in the univariate analysis (response variables which have a very low F value for subtotal would not be included in the multivariate run). When the multivariate run is complete the user may inspect the following results:

(a) Matrix of sums of squares and products based on error;

(b) Correlation matrix based on it;

(c) Likelihood - ratio test statistic for subtotal and attained probability level (recorded as 0, if it is < 10^{-9});

(d) Weights of the linear discriminant function (linear composite among response variables) which maximizes differences between cell effects; the largest characteristic root of $H(H+E)^{-1}$, with parameters ready for entry into Roy-Heck charts;

(e) Correlation of each response variable vs. the best discriminator. Steps (c), (d), and (e) are repeated for "Interaction", Row Effects, and Column Effects, with a modificaiton if the degrees of freedom for a given hypothesis is one (e.g., only two levels of rows or two levels of column effects). In this case, steps (c) and (d) are omitted, and only the F statistic (based on Hotelling's T^2) is reported with appropriate degrees of freedom.

There is a plotting facility (scatter plots for each selected pair of variables) very similar to the analysis-of-covariance unit (all data, a given row, column, or cell); of course there is no "regression function" plot, as this would be meaningless in pairs of random variables.

The most frequent use of this unit has been in multiple-stage analysis:

(a) Separate variables into non-discriminators and discriminators on the basis of the F value in univariate analysis; (b) separate the discriminators into classes - those that correlate highly (in absolute value) with the best discriminant function vs. those that do not; (c) select the latter for further discriminant analysis, especially if a factor has many levels, and for the analysis of all treatment effects (Subtotals). This process can be done, separately, for <u>all</u> effects, <u>row</u> effects, and <u>column</u> effects.

Novel features: Hierarchical discriminant analysis; reporting of likelihood ratio statistics (useful for testing) and union-intersection statistics (useful for confidence regions and variable selection)

Prior to its availability as an interactive module, this program was available in batch mode (MUDAID) on IBM and CDC computers. Since the number of response variables is often very large, the batch mode would

usually precede the conversational analysis. This unit was demonstrated (with transparencies) at Ames, Iowa, and in several meetings of the American Statistical Association. It has been used in the classroom (Multivariate Analysis) and for research studies, especially in educational research.

With the availability of very rapid turnaround time in batch mode in recent years, and because of the voluminous output of such multivariate analyses, the interactive version has had less use since 1975, with the very fast batch mode being widely used in and outside of the University of Georgia.

2.5. Product Flow Analysis

Documentation: Appendix R, Appendix vol. V, pp. 2-120, and THEMIS Report No. 27 (R. L. Wood [35]).

One of the popular management-information systems is the product-flow system, where a model describes the flow of materials from one station to the next; at each station, the product is processed by "servers" (usually machines or operators). A good first approximation to the queuing in such a production line is the M/M/s queue (exponential arrival, exponential departure, s servers) possibly with additional fixed delays at each station. Study of such product flow systems in an interactive mode are especially attractive to determine which queues will "explode" if servers in some stations are eliminated, and how many servers will be needed in each station if the process is to be scaled up to meet a specified demand(optimization of servers).

When the user calls $LINK PRODFLOW he receives instructions to declare options (Fig. 2.5.1).Typically, Code 2 in Option A would be chosen when a production line has been extensively studied in the pilot-plant stage and is ready for scale-up to production stage to meet a given demand. In Option B, certain service rate recommendations can also be sought - it is, of course, up to the production engineer to decide if such service rates are feasible. Option C determines display modes (whether the user wishes to have detailed reports on the probability of the length of each queue, or merely summary information of each station);

```
                              OUTPUT AREA

 IS PROGRAM IS DESIGNED AS A MANAGEMENT AID USING THE PROBABILITY
DISTRIBUTION APPROACH TO DETERMINE THE ATTRIBUTES OF A LINEAR MODEL.

THE MODEL CONTAINS FROM 1 TO 50 STATIONS, WHICH WILL BE DETERMINED
LATER.  EACH STATION HAS ASSOCIATED WITH IT THE NUMBER OF MACHINES,
MEAN SERVICE RATE, YIELD RATE, AND THE MAXIMUM ALLOWABLE WASTE.
    THE FOLLOWING ARE THE OPTIONS WHICH ARE AVAILABLE,  BY CODE.

  CODE      OPTION A
    1      USER SUPPLIES ALL DATA.
    2      USER SUPPLIES ALL DATA EXCEPT THE NUMBER OF MACHINES.

  CODE      OPTION B
    0      NO OPTIMIZATION.
    1      ALL  OPTIMIZATIONS.
    2      MACHINE OPTIMIZATION.
    3      SERVICE RATE RECOMENDATIONS

  CODE      OPTION C
    1      DISPLAY STATISTICS ONLY.
    2      DISPLAY ALL INFORMATION IN TABULAR FORM.
    3      DISPLAY ALL INFORMATION IN GRAPHIC FORM.

  CODE      OPTION D
    1      USER MAY SUPPLY THE MAXIMUM QUEUE LENGTHS.
    2      MAXIMUM QUEUE IS CALCULATED.

  A HARD COPY OF ANY DISPLAY MAY BE OBTAINED BY AIMING THE
LIGHT PEN AT THE SCREEN AND PRESSING THE PEDAL.  AFTER THIS RUN
MAKE A BATCH RUN TO GET THE TAPE.  THEN HAVE CALCOMP PLOT IT.
PRESS KEY 30 TO SET THE ARGUMENTS NECESSARY.
HARD COPIES MAY ALSO BE OBTAINED BY PRESSING KEY 16.

PRESS KEY 20 TO RE-START THIS PROGRAM AT ANY TIME.

PRESS KEY 1 TO PROCEED, IF SATISFIED.
PRESS KEY 31 TO TERMINATE THIS PROGRAM AT ANY TIME.

--------------------------------------------------------------------
                              REPLY AREA
```

Fig. 2.5.1

option D determines the handling of exploding queues - whether a certain percentage of "waste" is specified, or whether a maximum queue length is to be stated.

Figure 2.5.2 is an example of the echo to a user-supplied set of options - this run is typical of a first simulation of the line (no demand data, no optimization, all information user-supplied). Notice that for Station 4, a "tolerable" waste percentage must be specified, for when Station 3 works at full capacity (4 servers, each at a rate of 2.5 units per hour, with a yield of 90 percent) it supplies station 4 at a rate of 9 units per hour, which is exactly the maximum rate which Station 4 can handle (birth rate = death rate). Thus the queue can explode.

Figure 2.5.3 shows the summary information of the initial station which, by definition, works at full capacity. An example of a very detailed report on queue-length distributions is presented for Station 5, in Figure 2.5.4. Where (toward the end) probabilities of a given queue length (queue = numbers of units being served plus number of units awaiting service) are less than .01, queue lengths (26-29, 30 and greater) are lumped together. Figure 2.5.5 shows the same information in condensed pictorial form, and is usually all that is required in the day-to-day study of the simulated process line.

Figure 2.5.6 shows at which rate final product can be expected to arrive. Figure 2.5.7 is the first frame of an optimization run, in which the number of servers in each station is to be determined, so that a demand rate of 75 units can be met. It shows that 41 servers would be needed in Station 1. Where a pilot line works in a nearly optimal capacity, this is, of course, easy; but where several critical points occur in a system, the very simple queuing model can pinpoint the optimum number of servers more appropriately.

Novel features: Numerical and pictorial display of queue buildup in each station of a production-flow process. Application of queuing theory to prediction of service capacity to meet a specified demand rate of final product.

OUTPUT AREA

PRESS KEY 1 TO PROCEED, IF SATISFIED.
PRESS KEY 31 TO TERMINATE THIS PROGRAM AT ANY TIME.

IF YOU HAVE ENTERED YOUR DATA THROUGH A BATCH RUN, PRESS KEY 25.
IF YOU WISH TO RUN THE SAME DATA FROM YOUR LAST RUN, PRESS KEY 26.
IF YOU WISH TO ENTER NEW DATA, PRESS KEY 21.

HOW MANY STATIONS DO YOU HAVE?
5
WHAT IS THE DEMAND FOR THIS MODEL?
UNDER THE GIVEN OPTION, DEMAND IS NOT REQUIRED.
ENTER YOUR DATA AS FOLLOWS FOR EACH STATION.
PRESS KEY 10 TO RE-ENTER THE PREVIOUS STATION.

STATION NUMBER	NUMBER OF MACHINES	WITH AVERAGE SERVICE RATE	PERCENT YIELD	PERCENT WASTE	MAXIMUM QUEUE OPTIONAL
1	4.	3.00	90.00	0.0	0.
2	5.	2.20	90.00	0.0	0.
3	4.	2.50	90.00	0.0	0.
4	3.	3.00	95.00	20.00	0.
5	5.	2.00	90.00	0.0	0.

PRESS KEY 1 TO PROCEED, IF SATISFIED.
PRESS KEY 31 TO TERMINATE THIS PROGRAM AT ANY TIME.

REPLY AREA

Fig. 2.5.2

```
                              OUTPUT AREA

ENTER THE NUMBER OF THE STATION YOU WANT TO EXAMINE.
PRESS KEY 15 TO SEE THE FINAL PRODUCT.

          *** QUEUE DISTRIBUTION FOR STATION    1 ***

   **  MACHINES        4
   **  MEAN SERVICE RATE/MACHINE      3.00
   **  EFFECTIVE YIELD RATIO          0.90
   MAXIMUM QUEUE LENGTH ALLOWED AT THIS STATION IS  100
   PROBABILITY OF QUEUE EXPLODING AT THIS STATION IS  0.0
   TOLERABLE WASTE IS     25  PERCENT
            MAXIMUM ARRIVAL RATE IS        12.0000
            SERVICE CAPACITY IS            12.0000
ENTER THE NUMBER OF THE STATION YOU WANT TO EXAMINE.
PRESS KEY 15 TO SEE THE FINAL PRODUCT.
```

REPLY AREA

Fig. 2.5.3

OUTPUT AREA

*** QUEUE DISTRIBUTION FOR STATION 5 ***

```
** MACHINES         5
** MEAN SERVICE RATE/MACHINE    2.00
** EFFECTIVE YIELD RATIO        0.90
MAXIMUM QUEUE LENGTH ALLOWED AT THIS STATION IS   50
PROBABILITY OF QUEUE EXPLODING AT THIS STATION IS  0.0
TOLERABLE WASTE IS    25  PERCENT
        MAXIMUM ARRIVAL RATE IS        8.5500
        SERVICE CAPACITY IS           10.0000
```

QUEUE LENGTH FROM	TO	PROBABILITY	CUMULATIVE PROBABILITY
0	0	0.036319	0.036319
1	1	0.044879	0.081198
2	2	0.080779	0.161977
3	3	0.106454	0.268431
4	4	0.109649	0.378080
5	5	0.091982	0.470062
6	6	0.077803	0.547864
7	7	0.066089	0.613954
8	8	0.056273	0.670227
9	9	0.047985	0.718211
10	10	0.040954	0.759166
11	11	0.034975	0.794141
12	12	0.029880	0.824021
13	13	0.025534	0.849555
14	14	0.021824	0.871380
15	15	0.018656	0.890035
16	16	0.015948	0.905984
17	17	0.013634	0.919618
18	18	0.011656	0.931274
19	20	0.018486	0.949761
21	22	0.013513	0.963274
23	25	0.013771	0.977045
26	29	0.010688	0.987733
30	****	0.012261	

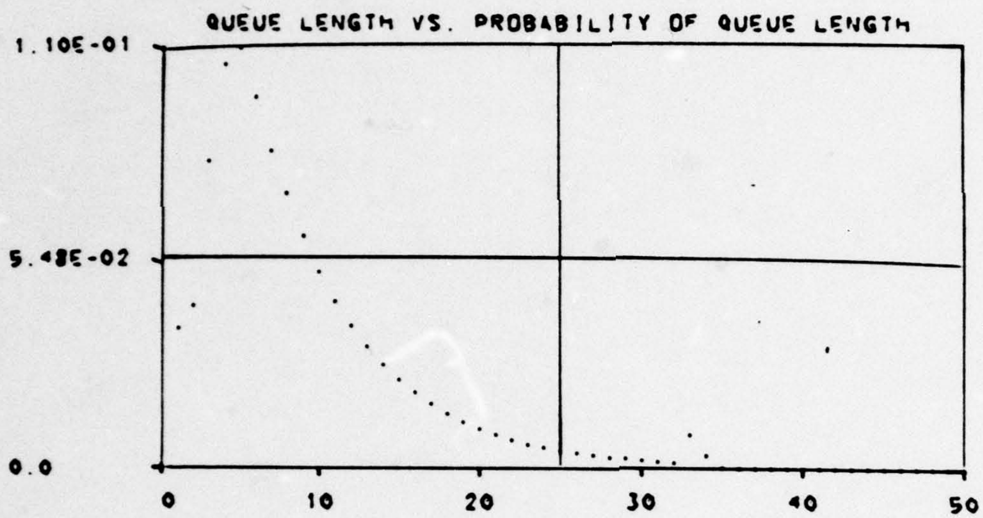
```
EXPECTED QUEUE LENGTH IS     7.711322
STANDARD DEVIATION IS        6.371940
```

PRESS KEY 1 TO PROCEED, IF SATISFIED.
PRESS KEY 31 TO TERMINATE THIS PROGRAM AT ANY TIME.

REPLY AREA

Fig. 2.5.4

```
*** QUEUE DISTRIBUTION FOR STATION    5 ***

**  MACHINES          5
**  MEAN SERVICE RATE/MACHINE      2.00
**  EFFECTIVE YIELD RATIO          0.90
MAXIMUM QUEUE LENGTH ALLOWED AT THIS STATION IS    50
PROBABILITY OF QUEUE EXPLODING AT THIS STATION IS  0.0
TOLERABLE WASTE IS     25  PERCENT
          MAXIMUM ARRIVAL RATE IS          8.5500
          SERVICE CAPACITY IS             10.0000
```




```
EXPECTED QUEUE LENGTH IS      7.771322
STANDARD DEVIATION IS         6.371940


PRESS KEY 1 TO PROCEED, IF SATISFIED.
PRESS KEY 31 TO TERMINATE THIS PROGRAM AT ANY TIME.
```

Fig. 2.5.5

```
                            OUTPUT AREA
ENTER THE NUMBER OF THE STATION YOU WANT TO EXAMINE.
PRESS KEY 15 TO SEE THE FINAL PRODUCT.

*** DISTRIBUTION RATE OF PRODUCTION OF FINAL PRODUCT ***

         FINAL PRODUCTS              PROBABILITY

              0.0                     0.036319
              1.80                    0.044879
              3.60                    0.080779
              5.40                    0.106454
              7.20                    0.109649
              9.00                    0.621920

        EXPECTED PRODUCTS OF               7.3332
        STANDARD DEVIATION OF              2.5653
        DEMAND FOR THIS MODEL WAS          0.0

PRESS KEY 1 TO PROCEED, IF SATISFIED.
PRESS KEY 31 TO TERMINATE THIS PROGRAM AT ANY TIME.
```

Fig. 2.5.6

```
                              OUTPUT AREA

 IF YOU HAVE ENTERED YOUR DATA THROUGH A BATCH RUN, PRESS KEY 25.
 IF YOU WISH TO RUN THE SAME DATA FROM YOUR LAST RUN, PRESS KEY 26.
 IF YOU WISH TO ENTER NEW DATA, PRESS KEY 27.
 PRESS KEY 26 IF YOU WISH TO CHANGE THE DEMAND REQUIRED.
 PRESS KEY 27 IF YOU WISH TO USE THE SAME DEMAND.
 WHAT IS THE DEMAND FOR THIS MODEL?
15

 THIS MODEL HAS     5  STATIONS.
 ONLY THE SERVICE RATE AND YIELD RATIOS ARE USER SUPPLIED.

 THE DEMAND TO BE SUPPLIED IS    15.00
 MACHINE OPTIMIZATION REQUESTED FOR THIS MODEL.

          *** INPUT DATA FOLLOWS ***

 STATION  NUMBER OF WITH AVERAGE PERCENT PERCENT MAXIMUM
 NUMBER   MACHINES  SERVICE RATE  YIELD   WASTE   QUEUE
    1                    3.00      0.90    0.25    100
    2                    2.20      0.90    0.25     50
    3                    2.50      0.90    0.25     50
    4                    3.00      0.95    0.20     50
    5                    2.00      0.90    0.25     50

                    FOR EXAMPLE.
         STATION 1 HAS   41 MACHINES
         EACH WITH A MEAN SERVICE RATE OF   3.00
         AND EFFECTIVE YIELD RATIO OF 0.90

 PRESS KEY 1 TO PROCEED, IF SATISFIED.
 PRESS KEY 31 TO TERMINATE THIS PROGRAM AT ANY TIME.
 PRESS KEY 5 TO RE-ENTER THESE VALUES.
 PRESS KEY 10 TO CHANGE SELECTED STATIONS.
```

Fig. 2.5.7

This unit was used in classroom instruction (Course on Information Systems, STA 804) and, in thesis research, to study estimation procedures in catenary compartmental systems. For the transient case a simulation option is available (since evaluation of convolutions of modified Bessel functions, needed for exact determination of probabilities, is very time-consuming). Adaptation of this, and most other, phases of this feature to a non-graphics system is, of course, very simple, but has not yet been done because of lack of demand at the University of Georgia.

2.6 Interactive OMNITAB

Documentation: Appendix V, appendix vol. VI, pp. 2-144, and THEMIS Report No. 31, (Bingham and Bargmann [4])

The central feature of OMNITAB (National Bureau of Standards) is a worksheet in which computations take place. Most of the commands of OMNITAB are column-oriented (operation performed on a specified column) or matrix manipulative, with matrices described by locating their beginning coordinates on the worksheet. With our ability to display large data matrices instantly, on a graphics terminal, it was appropriate to attempt implementation of an interactive version. The user of our interactive version needs to be familiar with the basic OMNITAB commands, and the positioning and meaning of operands. However, these are so easy to learn that, in our frequent application of this unit in the classroom, it took students only a few minutes to familiarize themselves with the most important commands (including matrix operations and statistical routines) to execute quite difficult problems.

In our interactive version we added, to the standard OMNITAB commands, the statistical distribution functions

YORMX, YORMP, YORMZ, BETAX, BETAP, BETAZ, GAMX, GAMP, GAMZ, TTX, TTP, TTZ, CHIX, CHIP, CHIZ, FFX, FFP, FFZ.

By using a modular facility which enables a programmer to add new commands to our system (described in Chapter IV, pp. 127 to 139 of THEMIS Report No. 31), we also added the matrix commands MTRI, MTRIN,

MCRROTS, MCVECT. The column-oriented distribution commands evaluate, for the given column, the normal distribution, beta distribution, gamma distribution, t-distribution, χ^2 - distribution, or F distribution; if the command ends in X, the probability is found, given an abscissa; if it ends in P, the abscissa ("percentage point") corresponding to an input probability is found, and if it ends in Z the ordinate for a given abscissa is found. As in all other column-oriented OMNITAB commands, the arguments can be fixed numbers (entered in real notation with decimal point) or column numbers (entered as integers).

The matrix commands MTRI obtains a lower triangular matrix T from a symmetric, positive-definite matrix Q, such that TT' = Q; if Q is singular, positive-semidefinite, the matrix T will be rectangular (n by r, where n is the order, r the rank of Q). MTRIN obtains the inverse of T, or an inverse from the left (of order r by n) if Q is singular. The matrix commands MCROOT and MCVECT obtain characteristic roots and vectors of symmetric matrices, a feature especially useful for multivariate statistical analysis. A list of OMNITAB commands available on our interactive system is displayed in Figure 2.6.1. This display is available to the user on request, by depressing Program Function Key 2. The unit was fully operational and transportable, on 370/158, under MVS. The attached figures are photographs of the screen.

Figure 2.6.2 has the record of commands used in a session performed during our October/November 1977 trials; such a record is available to the user, at any time, by pressing Program Function Key 2; he can also see a list of all OMNITAB commands currently on the system by depressing Program Function Key 3.

The central feature of our interactive version is the ability to display portions of the worksheet instantly, after each operation. Twelve lighted Program Function Keys are arranged in geometric analogy to the worksheet, each causing display of a 40 by 5 array (the total worksheet in our interactive mode is 80 by 30). (See Fig. 2.6.2 to 2.6.5)

The session is initiated by a call to $LINK OMTAB. A discussion of the enclosed illustration may help the reader to understand the operation of our version; ERASE is mandatory, since all our programs are

```
                              OUTPUT AREA
                  COMMANDS CURRENTLY IMPLEMENTED IN OMNITAB
[illegible]D ARCSIN   CHIP     EXPAND   M(K'AK)  MOVE     RESET
[illegible]RVEC ARCTAN CHIK    EXPONENT M(KAK')  MRAISE   RMS
ABS      ASCALAR  CHIZ     FFP      M(K'K)   MSCALAR  ROW
ABSOLUTE ASIN     CLOSE    FFK      M(KK')   MSUB     ROWSUM
ACOS     ASIND    COS      FFZ      MADD     MTRANS   SCALAR
ACOSD    ASINH    COSD     FLIP     MAX      MTRI     SET
ACOSH    ASUB     COSH     GAMP     MAXIMUM  MTRIN    SHORTEN
ACOT     ATAN     COT      GAMK     MCROOT   MULT     SIN
ACOTD    ATAND    COTD     GAMZ     MCVECT   MULTIPLY SIND
ACOTH    ATANH    COTH     GENERATE MDEFINE  MVECDIAG SINH
ADD      ATRANS   COUNT    HIERARCH MDIAG    MVECMAT  SORT
ADEFINE  AVECARR  DEFINE   INVERT   MERASE   MZERO    SQRT
ADIAG    AVECDIAG DEMOTE   LINEAR   MIDEN    NEGEXP   SUB
ADIVIDE  AVERAGE  DEVNOR   LOG      MIN      ORDER    SUBTRACT
AERASE   AZERO    DIV      LOGE     MINIMUM  PARPROD  SUM
AMOVE    BETAP    DIVIDE   LOGTEN   MINVERT  PARSUM   TAN
AMULT    BETAK    DUPLICAT M(AD)    MLINEAR  PRODUCT  TAND
ANTILOG  BETAZ    ERASE    M(AV)    MMATVEC  PROMOTE  TANH
ARAISE   BLOCKTRA EXCHANGE M(DA)    MMOVE    RAISE    TTP
ARCCOS   CHANGE   EXP      M(V'A)   MMULT    READ     TTK
ARCCOT
READY
```

REPLY AREA

Fig. 2.6.1

```
                         OUTPUT AREA
  THE SCREEN BECOMES FULL AN ALARM WILL SOUND.   WHEN YOU WANT TO SEE
 HE NEXT SECTION OF YOUR PROGRAM, PRESS KEY 2.

ERASE
SET 1
 001,.005,.01,.25,.5,.75,.90,.95,.975,.99,.995,.999
READ 1
ROW 4
 025
 05
 50
YORMP 1 2
CHIP 1 7. 3
TTP 1 27. 4
FFP 1 7 27 5
FFP 1 7. 27. 5
GENERATE -5..5.5..7
DEFINE 1..6
MULT 7 7 8
MULT 7 8 9
MULT 7 9 10
MI('() 1 6 21 5 23 6
MCVECT 23 6 4 4 29 1
MMOVE 29 1 4 4 29 6
MCROOT 23 6 4 4 29 10
MINVERT 23 6 5 5 34 6
MMOVE 38 10 1 1 1 11
RESET NRMAX 1
DIVIDE 1. 11 12
MMOVE 1 12 1 1 40 6
MSCALAR 34 10 4 1 440.64 3 11
MSCALAR 3 11 4 1 -1. 3 13
MMULT 1.1 10.4  3.13 4.1 10.11
MMULT 11.1. 11.4 3.13 4.1 20.11
MMOVE 1 5 21 1 10 12
MMULT 1 6 11 4 3 13 4 1 10 11
MMULT 12.6 10.4 3.13 4.1 21.11
MMOVE 1 10 21 1 10 12
MSUB 10 12 21 1 10 11 21 1 10 13
MI('() 10 13 21 1 10 15
READY
```

REPLY AREA

Fig. 2.6.2

overlayed into a small segment of core. SET 1 enters twelve probabilities into Worksheet Column 1. Three values were replaced with the READ 1 instruction (the echo following the READ command is ROW 1; if the user wishes to start elsewhere he types ROWX, and the program counts entries from this point on), (see detailed description on pp. 1-16 of THEMIS Report No. 31). Depressing Program Function Key 10 (the key corresponding to the top-left portion of the worksheet) the user would see only the first column of Worksheet Part 1, shown in Figure 2.6.3.

The command YORMP 1 2 in Fig. 2.6.2 obtains the percentage points under the normal distribution for the values in Col. 1 and stores them into Col. 2. The next command CHIP 1 7. 3 uses, again, the probability values of column 1 as input, the constant 7. as the degree of freedom (a real number, hence not a column number) and places percentage points, under the χ^2- distribution, with 7 degrees of freedom, into column 3. TTP 1 27. 4 obtains, in col. 4, the percentage points corresponding to col. 1, under the Student t distribution, with 27 degrees of freedom. FFP 1 7. 27. 5 (the first FFP call was erroneous, and the program gave the option to press Program Function Key 2 to ignore it - see THEMIS report No. 31) produces, in col. 5, the percentage points under the F-distribution, corresponding to the values in col. 1, with 7 and 27 df. Figure 2.6.3 is the output of all these operations - typically, this page of the worksheet was called after each command.

The remainder of this demonstration is an example of the kind of operations used frequently in classes of statistical methods and multivariate analysis - with the facility to see, instantly, the results of each matrix operation. A matrix of order 21 by 5 has been stored into columns 6 to 10. The OMNITAB command M(X'X) 1 6 21 5 23 6 takes this matrix ("the matrix starting at coordinate point 1, 6 of order 21 by 5"), performs the X'X multiplication, and stores the result into a field beginning in row 23, col. 6. MCVECT (an operation we added to our system) obtains the characteristic vectors of the 4 by 4 submatrix starting at (23, 6), and places them into the field beginning at (29, 1) (later

OUTPUT AREA

WORKSHEET PART 1

C O L U M N S

	1	2	3	4	5
1	0.100000E-02	-3.09023	0.598494	-3.42103	0.792369E-01
2	0.500000E-02	-2.57583	0.989256	-2.77068	0.131860
3	0.100000E-01	-2.32635	1.23904	-2.47266	0.165872
4	0.250000E-01	-1.95996	1.68987	-2.05183	0.228004
5	0.500000E-01	-1.64485	2.16735	-1.70329	0.294872
6	0.500000	0.0	6.34581	0.0	0.929523
7	0.900000	1.28155	12.0170	1.31370	1.95151
8	0.950000	1.64485	14.0671	1.70329	2.37321
9	0.975000	1.95996	16.0128	2.05183	2.80212
10	0.990000	2.32635	18.4753	2.47266	3.38822
11	0.995000	2.57582	20.2777	2.77068	3.85013
12	0.999000	3.09022	24.3218	3.42101	4.99823
13	0.0	0.0	0.0	0.0	0.0
14	0.0	0.0	0.0	0.0	0.0
15	0.0	0.0	0.0	0.0	0.0
16	0.0	0.0	0.0	0.0	0.0
17	0.0	0.0	0.0	0.0	0.0
18	0.0	0.0	0.0	0.0	0.0
19	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0
21	0.0	0.0	0.0	0.0	0.0
22	0.0	0.0	0.0	0.0	0.0
23	0.0	0.0	0.0	0.0	0.0
24	0.0	0.0	0.0	0.0	0.0
25	0.0	0.0	0.0	0.0	0.0
26	0.0	0.0	0.0	0.0	0.0
27	0.0	0.0	0.0	0.0	0.0
28	0.0	0.0	0.0	0.0	0.0
29	0.0	0.0	0.0	0.0	0.0
30	0.0	0.0	0.0	0.0	0.0
31	0.0	0.0	0.0	0.0	0.0
32	0.0	0.0	0.0	0.0	0.0
33	0.0	0.0	0.0	0.0	0.0
34	0.0	0.0	0.0	0.0	0.0
35	0.0	0.0	0.0	0.0	0.0
36	0.0	0.0	0.0	0.0	0.0
37	0.0	0.0	0.0	0.0	0.0
38	0.0	0.0	0.0	0.0	0.0
39	0.0	0.0	0.0	0.0	0.0
40	0.0	0.0	0.0	0.0	0.0

REPLY AREA

Fig. 2.6.3

moved to (29, 6) to have it, for convenience, in the same display as X'X). The four characteristic roots are on the far right (4 values beginning in 29, 10). The 5 by 5 matrix starting at 34, 6 contains the inverse of X'X. Its last element is moved to (1, 11) in order to have it in the next display. The setting of NRMAX to 1 insures that, until a new value is specified, all column commands will operate on the first number in a column only, but matrix commands are not affected.

The command MSCALAR appearing after the MMOVE shows another feature which makes OMNITAB so attractive: The scalar multiplier is reported as *40, 6* which means that this constant is to be taken from the worksheet, the number appearing at the (40,6) coordinate. Figures 2.6.4 and 2.6.5 show the results of all these operations; the pages of the worksheet were displayed after each operation.

Novel features: A facility to display segments of the OMNITAB worksheet *immediately* after each operation. A facility to add commands to the system, conforming with the general OMNITAB characteristics. Addition of statistical distribution functions to the OMNITAB commands, and thus the facility to construct tables instantly, especially for unusual values (e.g., fractional degrees of freedom or extreme probability values).

Since its availability in 1972, this has been the most popular unit. Some of the designers of OMNITAB saw this unit in operation. It was reported at the national meetings of the American Statistical Association in New York (1973) and Atlanta (1975). Students in the Multivariate Analysis classes quickly learned how to operate this unit and could, step-by-step, perform canonical-partial correlation problems and factor analysis problems in a very short time (15 to 30 minute sessions for each student).

It is only since the advent of the interactive SAS matrix manipulations at our campus (late 1976) that there has been anything comparable to this facility. Adaptation to more modern equipment has not yet been undertaken, since large instant display is a very important feature, and teletype terminals may not be fast enough to eliminate the element of user fatigue (see discussion of the Analysis of Covariance adaptation in Section 3.3).

OUTPUT AREA

WORKSHEET PART- 2

COLUMNS

	6	7	8	9	10
1	1.00000	-5.00000	25.0000	-125.000	625.000
2	1.00000	-4.50000	20.2500	-91.1250	410.063
3	1.00000	-4.00000	16.0000	-64.0000	256.000
4	1.00000	-3.50000	12.2500	-42.8750	150.063
5	1.00000	-3.00000	9.00000	-27.0000	81.0000
6	1.00000	-2.50000	6.25000	-15.6250	39.0625
7	1.00000	-2.00000	4.00000	-8.00000	16.0000
8	1.00000	-1.50000	2.25000	-3.37500	5.06250
9	1.00000	-1.00000	1.00000	-1.00000	1.00000
10	1.00000	-0.500000	0.250000	-0.125000	0.625000E-01
11	1.00000	0.0	0.0	0.0	0.0
12	1.00000	0.500000	0.250000	0.125000	0.625000E-01
13	1.00000	1.00000	1.00000	1.00000	1.00000
14	1.00000	1.50000	2.25000	3.37500	5.06250
15	1.00000	2.00000	4.00000	8.00000	16.0000
16	1.00000	2.50000	6.25000	15.6250	39.0625
17	1.00000	3.00000	9.00000	27.0000	81.0000
18	1.00000	3.50000	12.2500	42.8750	150.063
19	1.00000	4.00000	16.0000	64.0000	256.000
20	1.00000	4.50000	20.2500	91.1250	410.063
21	1.00000	5.00000	25.0000	125.000	625.000
22	0.0	0.0	0.0	0.0	0.0
23	21.0000	0.0	192.500	0.0	3166.63
24	0.0	192.500	0.0	3166.63	0.0
25	192.500	0.0	3166.63	0.0	61825.2
26	0.0	3166.63	0.0	61825.2	0.312500
27	3166.63	0.0	61825.2	0.312500	0.131040E 07
28	0.0	0.0	0.0	0.0	0.0
29	0.608556E-01	0.998147	0.0	0.0	3178.36
30	0.0	0.0	0.511169E-01	-0.998690	9.26354
31	0.998147	-0.608556E-01	0.0	0.0	61981.4
32	0.0	0.0	0.998690	0.511169E-01	30.2292
33	0.0	0.0	0.0	0.0	0.0
34	0.169233	0.511822E-06	-0.292099E-01	-0.311138E-01	0.969175E-03
35	0.511822E-06	0.329940E-01	-0.188128E-06	-0.168992E-02	0.804210E-08
36	-0.292099E-01	-0.188128E-06	0.904659E-02	0.114363E-01	-0.356234E-03
37	-0.311138E-01	-0.168992E-02	0.114363E-01	0.102731E-03	-0.488881E-09
38	0.969175E-03	0.804210E-08	-0.356234E-03	-0.488881E-09	0.152283E-04
39	0.0	0.0	0.0	0.0	0.0
40	65661.0	0.0	0.0	0.0	0.0

REPLY AREA

Fig. 2.6.4

OUTPUT AREA

WORKSHEET PART 3

C O L U M N S

	11	12	13	14	15
1	0.152283E-04	65667.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0
3	63.6428	0.0	-63.6428	0.0	0.0
4	0.528100E-03	0.0	-0.528100E-03	0.0	0.0
5	-23.3928	0.0	23.3928	0.0	0.0
6	-0.321034E-04	0.0	0.321034E-04	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0
10	521.176	625.000	103.824	0.0	65667.1
11	410.061	410.063	0.122070E-02	0.0	0.0
12	310.642	256.000	-54.6423	0.0	0.0
13	222.920	150.063	-72.8514	0.0	0.0
14	146.893	81.0000	-65.8934	0.0	0.0
15	82.5632	39.0625	-43.5001	0.0	0.0
16	29.9293	16.0000	-13.9293	0.0	0.0
17	-11.0083	5.06250	16.0708	0.0	0.0
18	-40.2495	1.00000	41.2495	0.0	0.0
19	-51.7943	0.625000E-01	51.8568	0.0	0.0
20	-63.6428	0.0	63.6428	0.0	0.0
21	-51.7948	0.625000E-01	51.8573	0.0	0.0
22	-40.2505	1.00000	41.2505	0.0	0.0
23	-11.0096	5.06250	16.0721	0.0	0.0
24	29.9277	16.0000	-13.9277	0.0	0.0
25	82.5616	39.0625	-43.4991	0.0	0.0
26	146.892	81.0000	-65.8920	0.0	0.0
27	222.919	150.063	-72.8565	0.0	0.0
28	310.642	256.000	-54.6423	0.0	0.0
29	410.062	410.063	0.244141E-03	0.0	0.0
30	521.179	625.000	103.821	0.0	0.0
31	0.0	0.0	0.0	0.0	0.0
32	0.0	0.0	0.0	0.0	0.0
33	0.0	0.0	0.0	0.0	0.0
34	0.0	0.0	0.0	0.0	0.0
35	0.0	0.0	0.0	0.0	0.0
36	0.0	0.0	0.0	0.0	0.0
37	0.0	0.0	0.0	0.0	0.0
38	0.0	0.0	0.0	0.0	0.0
39	0.0	0.0	0.0	0.0	0.0
40	0.0	0.0	0.0	0.0	0.0

REPLY AREA

Fig. 2.6.5

2.7. Response Surface Fitting

Documentation: THEMIS Report No. 37 (W. P. Bond and R. E. Bargmann [43]) pp. 40-78, 105-176, 255-314.

This interactive unit is designed to permit study of response surfaces, for up to four factors (variables) each with either 2, 3, or 4 ordinal levels*. The number of replications for each level combination (cell) must be equal. The surface fitted is thus a tensor-product of polynomials, up to the third degree in each of the 2, 3, or 4 variables.

The user calls $LINK PATS and is asked to define and name variables, encode levels (if unequally spaced) and enter data. Figure 2.7.1 shows the photograph of the screen when data entry is complete. The user may review the analysis of variance table (part of which is shown in Fig. 2.7.2) and the sum of squares (= mean square) associated with each orthogonal contrast (Fig. 2.7.3 - the coefficients are those of orthogonal polynomials of the response equation, e.g., the last (2,2) coefficient is that of $(3a^2-2)\cdot(3b^2-2)$, where a = 1 for high level, 0 for middle level, and -1 for low level of Factor A).

The user may now choose to eliminate some of the degrees-of-freedom, e.g., when F-ratios are too small for that contrast, or when the degree of the corresponding product of polynomials is too large. He may also wish to edit data.

After such modification, if any, the user has various options of viewing the response equation (Fig. 2.7.4). He may plot the response equation vs. the levels of one factor, with the other factor or factors specified at a given point, which need not be one of the exact levels (see Fig. 2.7.5, where FACB is set at 2, and 2.7.6 where FACB is set at 1); 95% confidence bounds are displayed around the response equation. This mode of display is especially useful for process evaluation (see pp. 143-165 of THEMIS report No. 37).

Another display mode (Fig. 2.7.7) enables the user to view contours of the response surface, for two factors, where the third and fourth factors (if any) are fixed at a specified level. Fig. 2.7.8 shows the

*If levels are nominal, the analyses of variance are still valid; the response equations are, of course, useless in that event.

```
                                              OUTPUT ARE:
  : FACA            B = FACB

      DEPRESS KEY 30 TO EDIT A LINE OF DA.

 CELL A B

  1   1 1      10,12,11
  2   2 1      21,22,23
  3   3 1      15,16,14
  4   1 2      18,20,18
  5   2 2      27,25,26
  6   3 2      22,20,20
  7   1 3      15,15,16
  8   2 3      26,26,25
  9   3 3      18,15,16
DATA ENTRY IS COMPLETE.
DEPRESS KEY 1 TO PERFORM ANALYSIS
DEPRESS KEY 2 TO REVIEW OR EDIT DATA.
```

Fig. 2.7.1

OUTPUT AREA

ANALYSIS OF VARIANCE

SOURCE OF VARIATION	DF	SUM OF SQUARES
FACA A	2	446.741
FACB A	2	150.518
FACA AFACB A	4	20.3704
ALL EFFECTS	8	617.629
ERROR	18	19.3333

Fig. 2.7.2

OUTPUT AREA

SINGLE DEGREE OF FREEDOM DISPLAY

DEPRESS KEY 1 TO DELETE TERMS BY CELL NUMBER
DEPRESS KEY 2 TO DELETE TERMS BY F COMPARISON
DEPRESS KEY 28 TO PAGE BACKWARD
DEPRESS KEY 29 TO PAGE FORWARD

DEPRESS KEY 30 TO VIEW OPTION TABLE

MSE = 1.0740738

A = FACA B = FACB

CELL	A	B	D	COEFFICIENT	SUM OF SQUARES	F RATIO
1	0	0		18.962952	9709.0352	9039.4453
2	1	0		1.1666660	24.500000	22.810349
3	2	0		-2.7962961	422.24072	393.12061
4	0	1		1.5555553	43.555542	40.551712
5	1	1		-0.75000000	6.7500000	6.2844839
6	2	1		-0.13888884	0.69444442	0.64655185
7	0	2		-1.4074068	106.96295	99.586212
8	1	2		0.83333313E-01	0.25000000	0.23275864
9	2	2		-0.34259254	12.675925	11.801726

REPLY AREA

Fig. 2.7.3

OUTPUT AREA

OPTION TABLE

```
DEPRESS KEY 1 TO PLOT THE PREDICTED RESPONSE AGAINST ONE FACTOR
DEPRESS KEY 2 TO PLOT A 95% CONFIDENCE INTERVAL AGAINST ONE FACTOR
DEPRESS KEY 3 TO PLOT CONTOURS
DEPRESS KEY 4 TO EDIT THE RAW DATA
DEPRESS KEY 5 TO RETURN TO THE RESPONSE EQUATION
```

Fig. 2.7.4

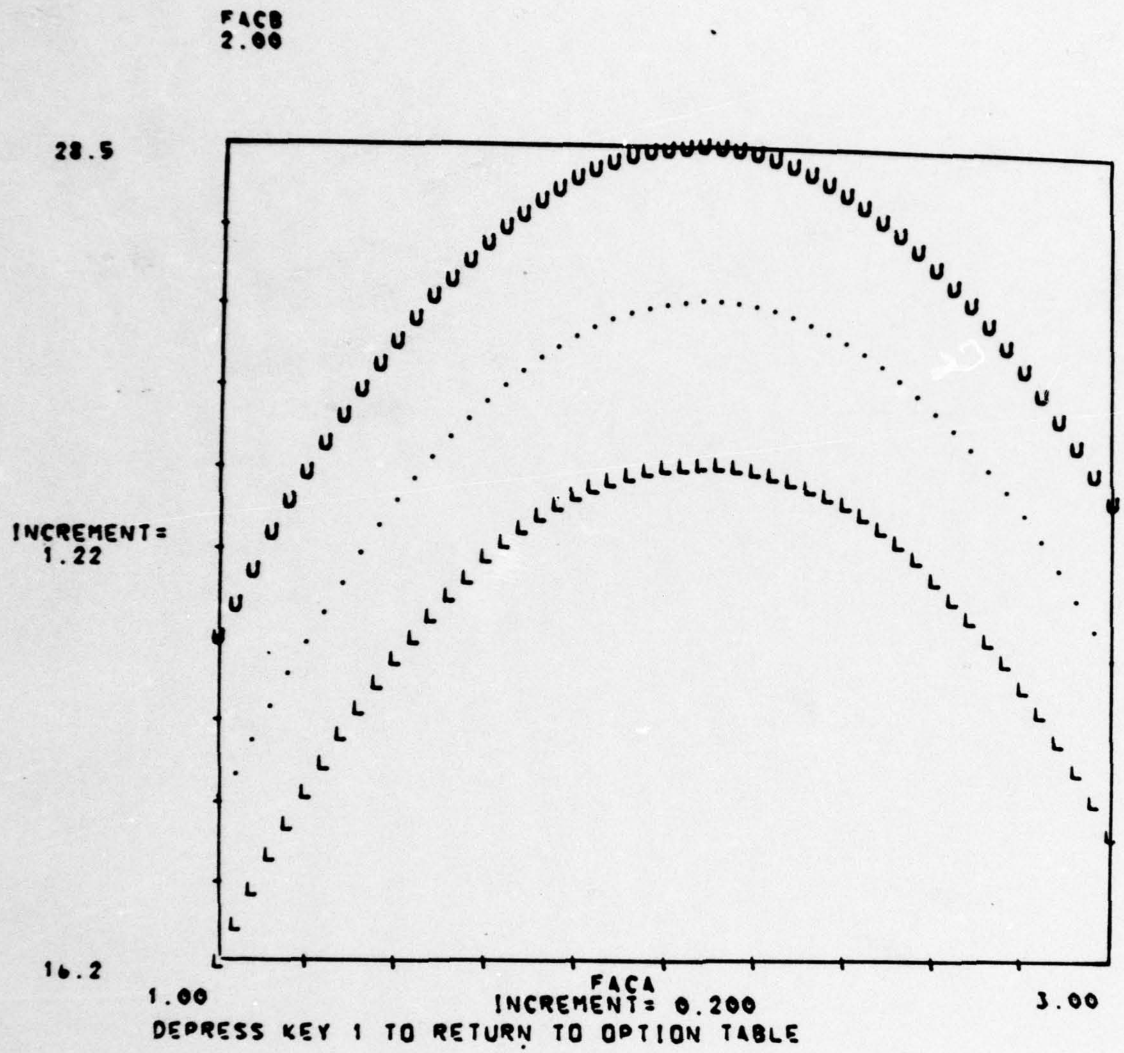

FACB
2.00
28.5
INCREMENT=
1.22
16.2
1.00
FACA
INCREMENT= 0.200
3.00
DEPRESS KEY 1 TO RETURN TO OPTION TABLE

Fig. 2.7.5

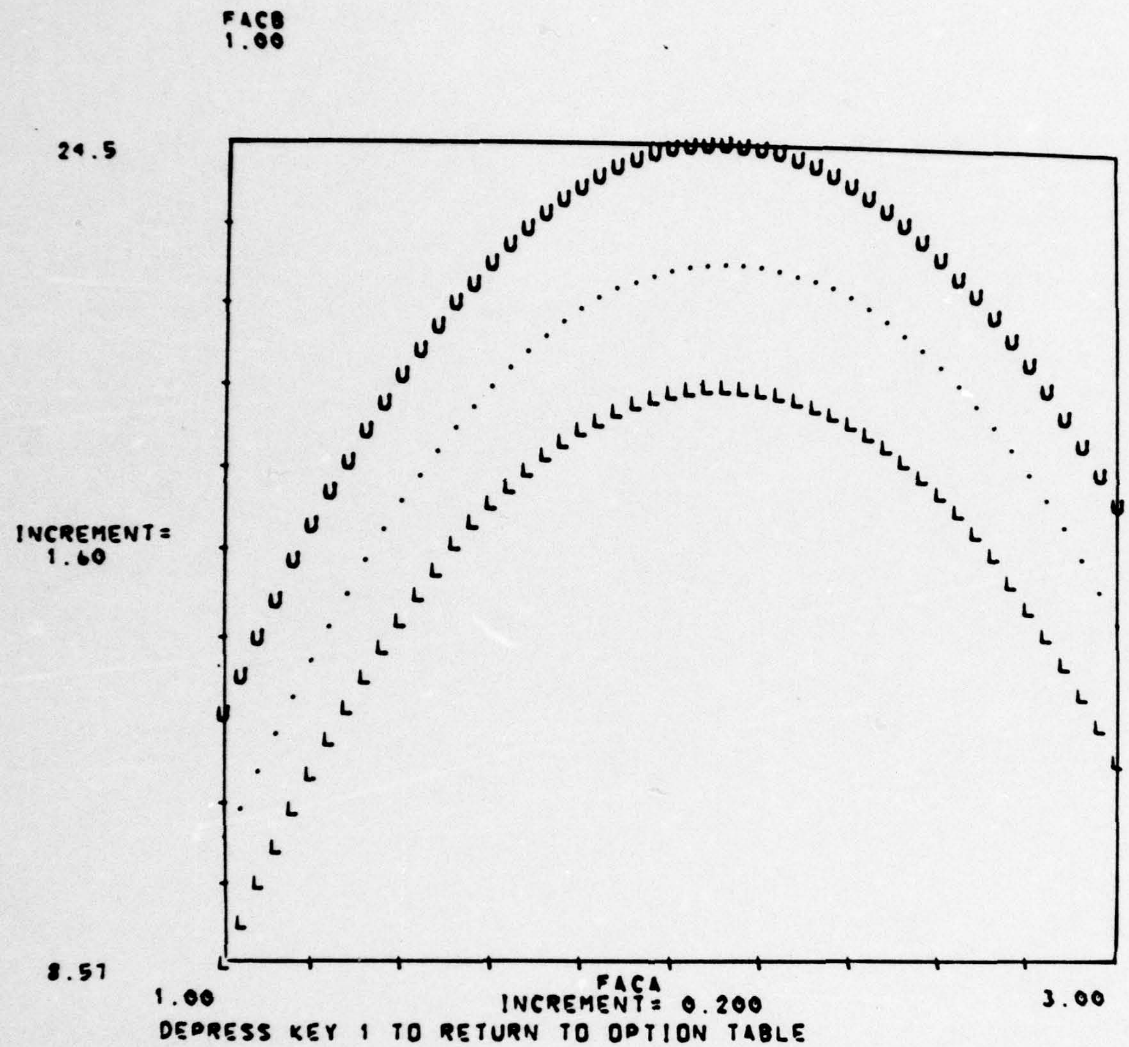

FACB
1.00
24.5
INCREMENT=
1.60
8.51
1.00
FACA
INCREMENT= 0.200
3.00
DEPRESS KEY 1 TO RETURN TO OPTION TABLE

Fig. 2.7.6

```
FACTORS ARE AS FOLLOWS:

1  FACA     1.3
2  FACB     1.3

ENTER, SEPARATED BY A COMMA, THE MINIMUM AND MAXIMUM
VALUE FOR THE INDICATED FACTOR

ENTER THE PREDICTED RESPONSE FOR WHICH YOU WISH TO PLOT A CONTOUR
```

Fig. 2.7.7

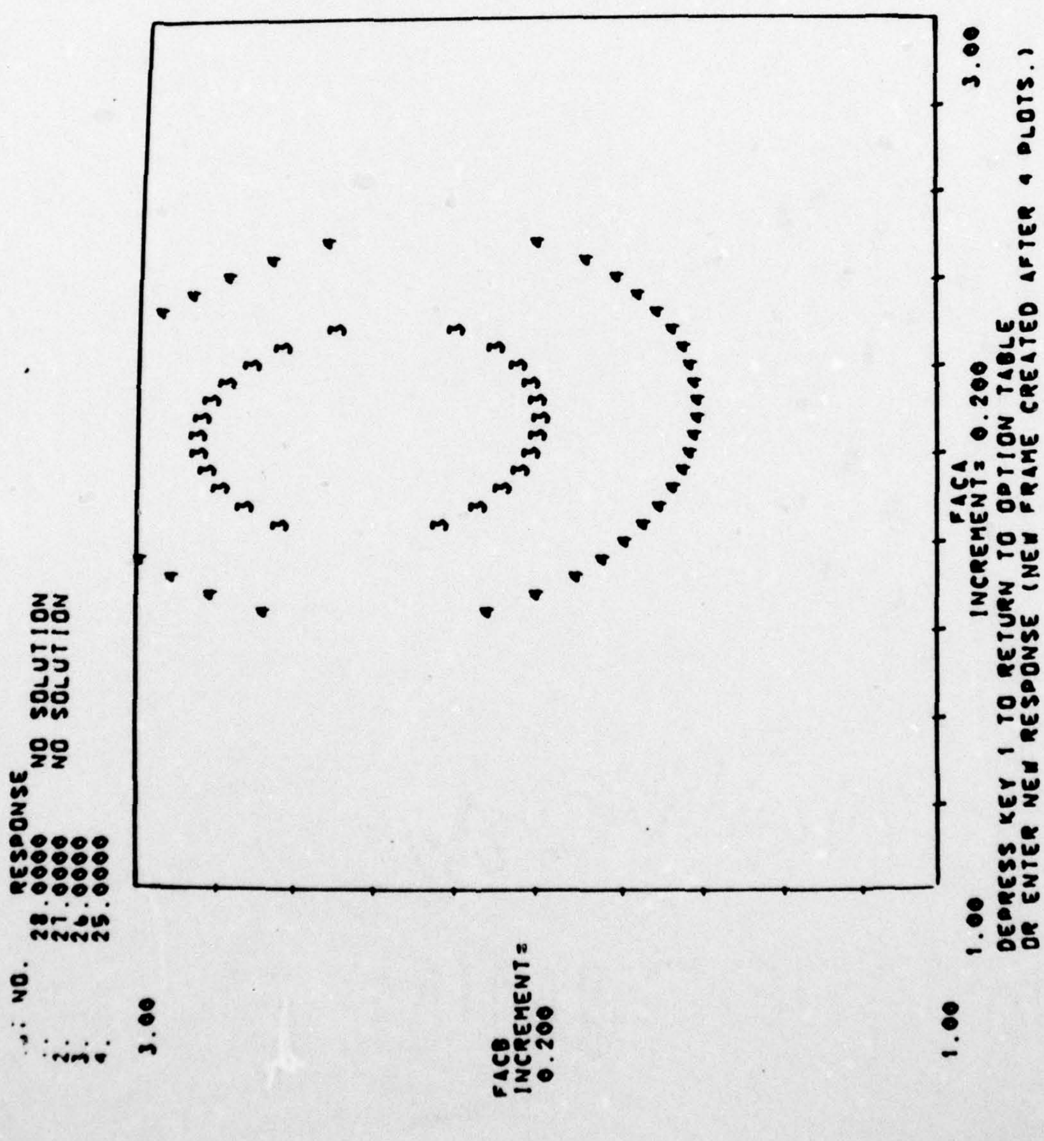

NO. RESPONSE
1: 28.0000 NO SOLUTION
2: 27.0000 NO SOLUTION
3: 26.0000
4: 25.0000
3.00
FACB
INCREMENT=
0.200
1.00
1.00
3.00
FACA
INCREMENT= 0.200
DEPRESS KEY 1 TO RETURN TO OPTION TABLE
OR ENTER NEW RESPONSE (NEW FRAME CREATED AFTER 4 PLOTS.)

Fig. 2.7.8

contour where the response is 25 (plotted as 4) and 26 (plotted as 3); the desired level of the response variable was entered by the user. The attainment of a maximum is shown in Fig. 2.7.9 (at 26.4 for the response) Figure 2.7.10 shows the algebraic expression of the response equation. This mode of display is useful for locating extrema, ridges, or saddle points of response surfaces, and for studying the behavior of likelihood functions involving two or more parameters (pp. 166-176 of THEMIS report No. 37).

Novel features: Efficient algorithms (extended Yates) for obtaining response equations. Plotting of projections of response equations and confidence regions. Calculations and display of contours for all levels of two factors, with the others held fixed. Facility to eliminate individual orthogonal contrasts.

This unit was operational on the 370/158, under MVS, in November, 1977. The figures shown here are photographs of the screen from these trials (the displays shown in THEMIS report No. 37 were CALCOMP plots of the screen display). Since the unit was not fully operational until two years after the end of the contract period, its use has been limited to classroom instruction in a course on Information Systems, and in a course on Non-linear Statistical Analysis.

2.8. Function Plotting Facility

Documentation: THEMIS Report No. 37 (Bond and Bargmann [43]) pp. 29-104, 177-231, 315,336.

This interactive unit is designed to prepare function plots, especially statistical distribution functions. It can plot simple algebraic expressions involving functions in the graphics calculator unit (see Section 3.2). Since it is designed especially for functions the evaluation of which takes considerable computer time, the exact evaluation takes place on only 11 points chosen over the domain of the abscissa specified by the user; the rest of the graph is obtained by fitting a cubic spline through these 11 points. In the region of final interest the precision of the function plot can be improved by the "blowup" technique, which fits the cubic spline for a range of X-values which can be narrowed at each stage.

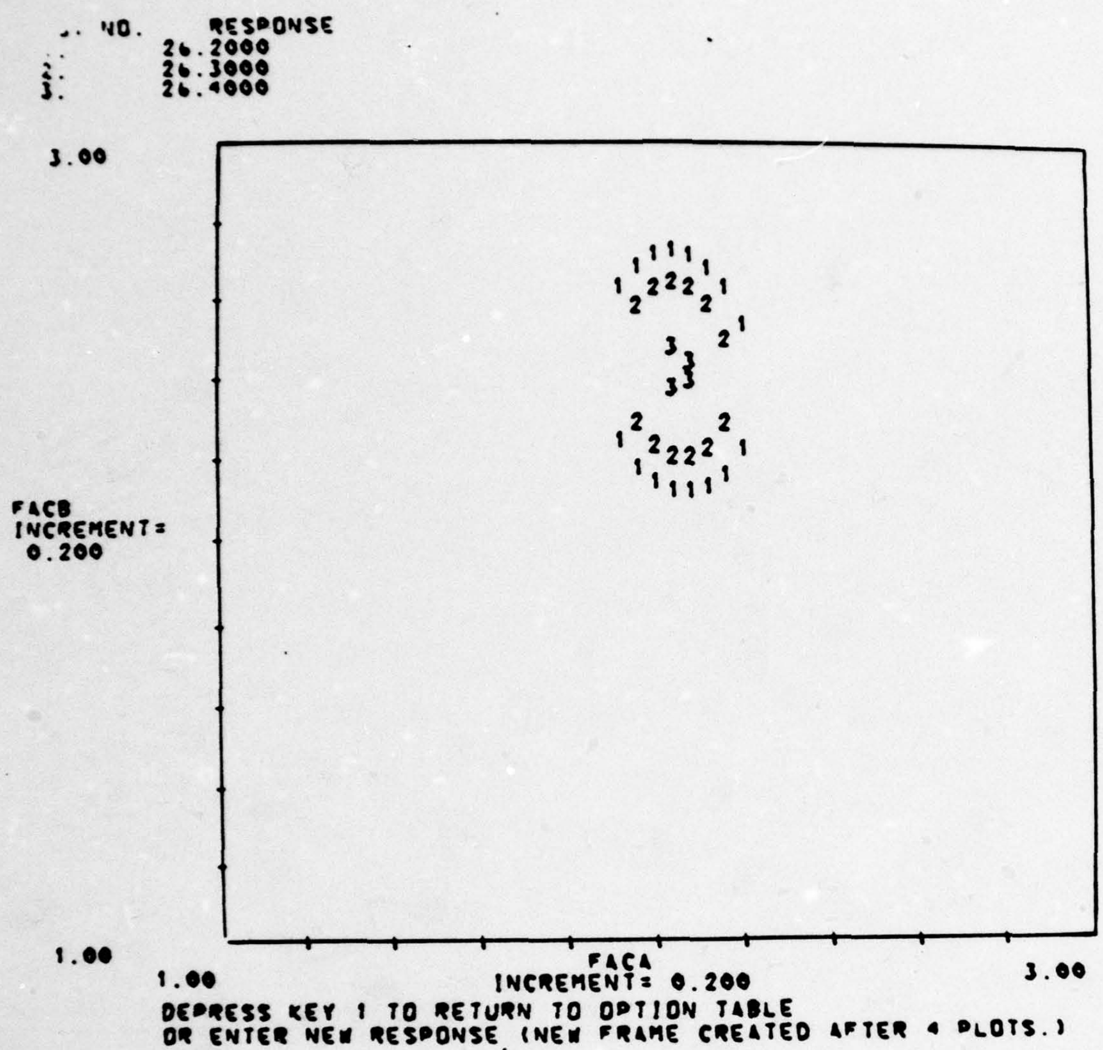

NO. RESPONSE
26.2000
2. 26.3000
3. 26.4000
3.00
FACB
INCREMENT=
0.200
1.00
1.00
FACA
INCREMENT= 0.200
3.00
DEPRESS KEY 1 TO RETURN TO OPTION TABLE
OR ENTER NEW RESPONSE (NEW FRAME CREATED AFTER 4 PLOTS.)

Fig. 2.7.9

```
                                    OUTPUT AREA
                        FACTOR*A    IS  FACA
                        FACTOR*B    IS  FACB

            IN STANDARD FACTOR VALUES, Y =
 +  25.99998        + 0.9999994      *A
  -6.333335     *A**2   +  1.833333      *B
 -0.7500000     *A*B  -0.4166665     *A**2 *B
  -2.166665     *B**2   + 0.2499999      *A*B**2
  -3.083330     *A**2 *B**2

     WHERE*A     = 0.0              * ARAW**2   + (   1.000000     )* ARAW
 + ( -2.000000      )
     WHERE*B     = 0.0              * BRAW**2   + (   1.000000     )* BRAW
 + ( -2.000000      )
     MINIMUM TWO SIGMA ERROR IS + OR -          2.11078
     MAXIMUM TWO SIGMA ERROR IS + OR -          2.42642




DEPRESS KEY 28 TO PAGE BACKWARD.
DEPRESS KEY 29 TO PAGE FORWARD.

DEPRESS KEY 30 TO VIEW OPTION TABLE
____________________________________________________________
                              REPLY AREA
```

Fig. 2.7.10

After calling $LINK BLOWUP the user receives instructions such as those exhibited in Figure 2.8.1. In this example, a variable parameter was introduced, called DF, its value set at 15, 30, 60, and 120. The function is to be plotted for values of X, in increments of .02, from -4 to +4 (note that, no matter what the desired increments are, only 11 points are determined precisely from the function subprograms). The function stated in the end (illegible in the photograph) was

Y = YORMZ(X) - TTZ2(X,DF)

denoting difference between the probability density functions of the normal distribution and t distributions with selected degrees of freedom. Figure 2.8.2 is an echo to the user-supplied input, and enables the user, with the help of program function keys, to change any line requested. Figure 2.8.3 is the plot for this input. Notice that one of the parameters, the minimum and maximum value of the Y-axis, can be changed at this stage. This facility was introduced because, frequently, the range of function values must be guessed at the first attempt and usually produces a poor plot (one poor plot preceded the one shown in Fig. 2.8.3).

The displays shown in Fig. 2.8.4 and 2.8.5 illustrate another important feature of this unit. The function POWER was not a member in the graphics calculator, and was included by a simple technique described on pp. 89-91, and illustrated on pp. 224-231 of THEMIS report No. 37. The built-in function was BETNCS (see Section 4.2) which evaluates the non-central beta distribution, but arguments for the latter need to be obtained by a call to BETAP first. The plots in Fig. 2.8.5 are power functions for the analysis-of-variance test, with 4 and 36 degrees of freedom, for α = .05, .01, and .001, and for the noncentrality parameter (γ^2/n) extending from 0 to 2.

Figures 2.8.6 and 2.8.7 illustrate the use of the "blowup" facility. It was desired, for the α = .01 case, to find that value of the non-centrality parameter for which the power would be .90; Fig. 2.8.6 shows that it is between .6 and .65; Fig. 2.8.7 provides further resolution and shows the desired value to be very close to .610.

```
                         OUTPUT AREA
THIS UNIT WILL PLOT ANY EXPRESSION CURRENTLY AVAILABLE
IN THE CALCULATOR MODE

THE FUNCTION CAN BE OF THE FORM

     Y = F(X) ,  OR
     Y = F(Z)  WHERE  Z = G(X)

IN ADDITION, THE EXPRESSION MAY BE PLOTTED FOR UP
TO FOUR VALUES OF A PARAMETER, SAY P.

DEPRESS KEY 31 TO TERMINATE THE SESSION.

IF YOU WISH TO SPECIFY A VARIABLE PARAMETER
DEPRESS KEY 2; OTHERWISE, DEPRESS KEY 1 TO CONTINUE

ENTER THE NAME BY WHICH YOU WILL
REFER TO THE PARAMETER

ENTER, SEPARATED BY COMMAS, UP TO FOUR VALUES FOR
OF
15,30,60,120

IF YOU WISH TO SPECIFY Z AS A FUNCTION OF X
DEPRESS KEY 2; OTHERWISE, DEPRESS KEY 1 TO CONTINUE

ENTER, SEPARATED BY COMMAS, THE MINIMUM AND MAXIMUM VALUES FOR X
-4,4
ENTER THE INCREMENT FOR X.
 02

ENTER, SEPARATED BY COMMAS, THE MINIMUM AND MAXIMUM VALUES FOR Y
- 05, 05

ENTER THE FUNCTION IN THE FORM  Y = F(X)
OR  Y = F(Z)  IF Z IS DEFINED

                          REPLY AREA
[illegible]_
```

Fig. 2.8.1

AD-A049 707 GEORGIA UNIV ATHENS DEPT OF STATISTICS AND COMPUTER--ETC F/G 12/1
INTERACTIVE STATISTICAL SOFTWARE.(U)
JAN 78 R E BARGMANN, H BOUVER N00014-69-A-0423
UNCLASSIFIED THEMIS-UGA-33 NL

2 OF 2
AD
A049707

END
DATE
FILMED
3 -78
DDC

```
                                     OUTPUT AREA
           SUMMARY
 1  Y=YGRMZ(X)-TTZ2(X,DF)
 2  NO Z = G(X) SPECIFIED
 3  MIN FOR X = -0.40000000D 01 MAX FOR X =  0.40000000D 01        Y
 4  INCREMENT FOR X =  0.20000000D-01
 5  MIN FOR Y = -0.50000000D-01 MAX FOR Y =  0.50000000D-01
 6  THE VALUES FOR DF        ARE:
      0.15000000D 02    0.30000000D 02    0.60000000D 02    0.12000000D 03
DEPRESS THE KEY CORRESPONDING TO ANY ITEM YOU WISH TO CHANGE
OR DEPRESS KEY 28 TO PLOT THE FUNCTION
```

REPLY AREA

Fig. 2.8.2

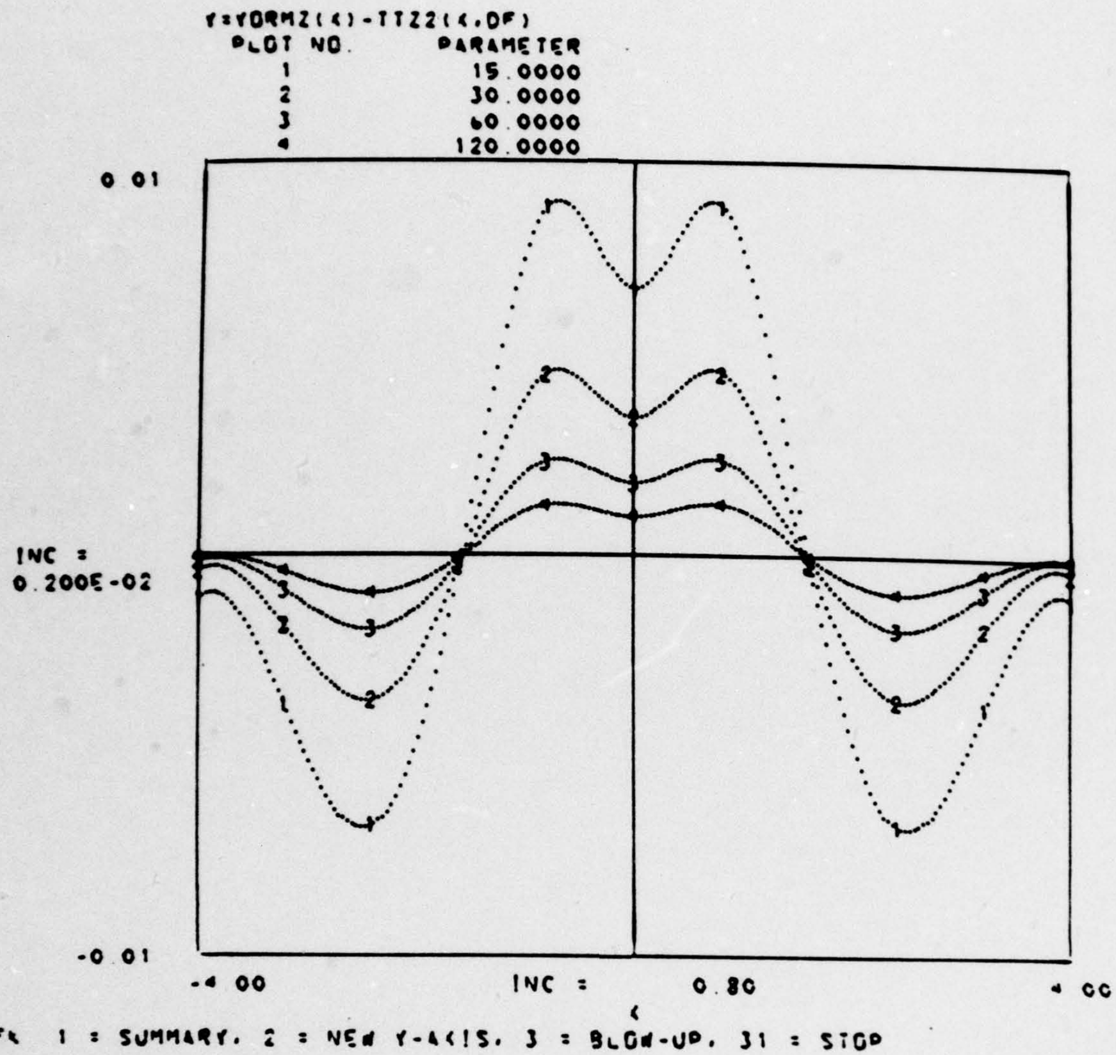
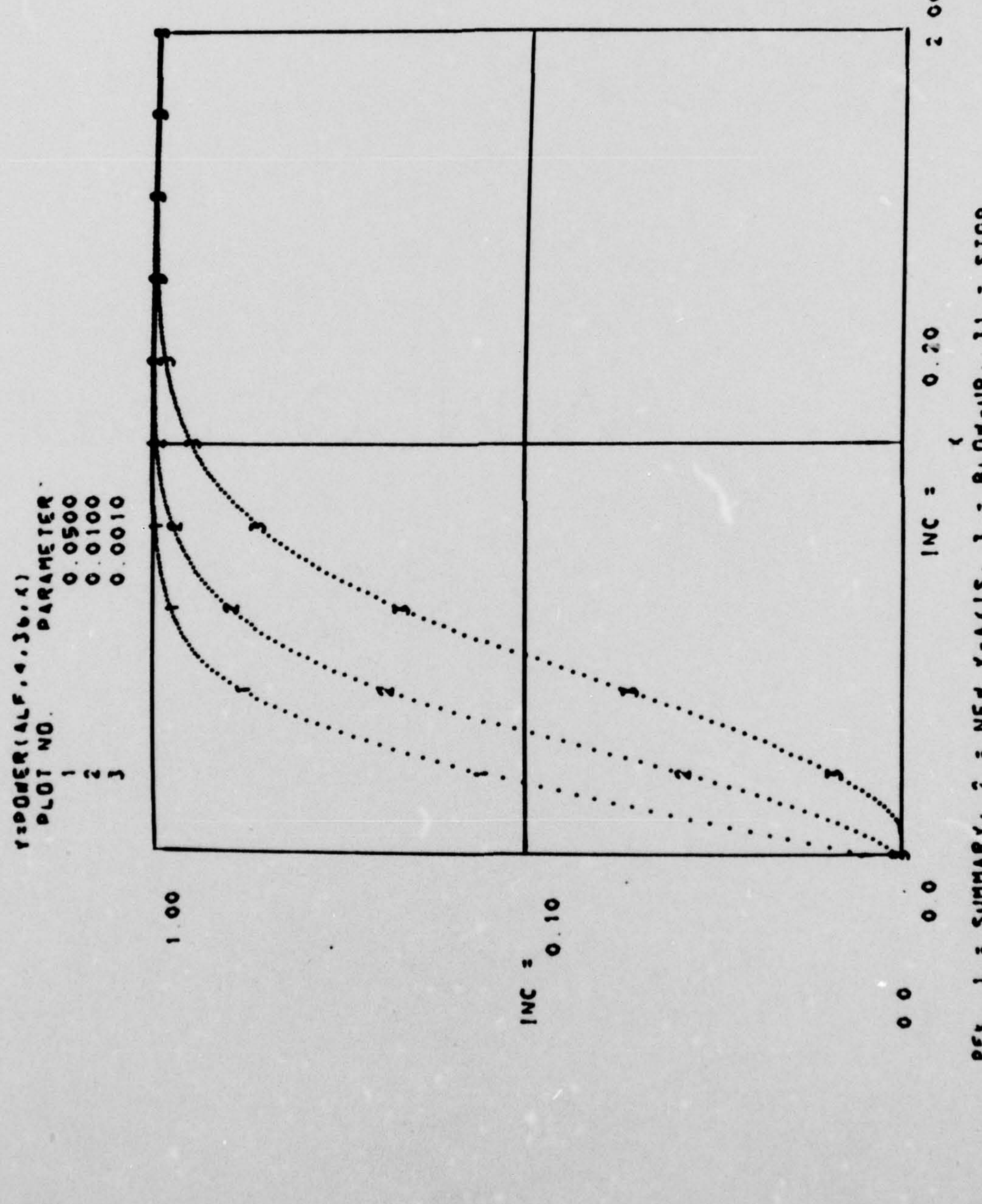

Y=YORMZ(K)-TTZ2(K,DF)
PLOT NO. PARAMETER
1 15.0000
2 30.0000
3 60.0000
4 120.0000
0.01
INC =
0.200E-02
-0.01
-4.00
INC =
0.80
4.00
K
OPT 1 = SUMMARY, 2 = NEW Y-AXIS, 3 = BLOW-UP, 31 = STOP

Fig. 2.8.3

```
                              OUTPUT AREA
           SUMMARY

1  Y=POWER(ALF,4,36,X)
2  NO Z = G(X) SPECIFIED
3  MIN FOR X =  0.0                 MAX FOR X =  0.2000000D 01
4  INCREMENT FOR X =  0.5000000D-02
5  MIN FOR Y =  0.0                 MAX FOR Y =  0.1000000D 01
6  THE VALUES FOR ALF         ARE:
      0.5000000D-01    0.1000000D-01    0.1000000D-02

DEPRESS THE KEY CORRESPONDING TO ANY ITEM YOU WISH TO CHANGE

OR DEPRESS KEY 28 TO PLOT THE FUNCTION
```

REPLY AREA

Fig. 2.8.4



Y=POWER(ALF,4,36,X)
PLOT NO. PARAMETER
1 0.0500
2 0.0100
3 0.0010
1.00
INC = 0.10
0.0
0.0
INC = 0.20
2.00
OPT 1 = SUMMARY, 2 = NEW Y-AXIS, 3 = BLOW-UP, 31 = STOP

Fig. 2.8.5

OUTPUT AREA

:POWER:LLF,4,36,K)

PLOT NO	PARAMETER
1	0.0500
2	0.0100
3	0.0010

0.40000	0.87626	0.67863	0.36451
0.45000	0.91793	0.75091	0.44192
0.50000	0.94580	0.80973	0.51838
0.55000	0.96358	0.85683	0.59155
0.60000	0.97498	0.89392	0.65900
0.65000	0.98315	0.92265	0.71880
0.70000	0.98908	0.94439	0.77088
0.75000	0.99318	0.96047	0.81562
0.80000	0.99590	0.97220	0.85342
0.85000	0.99764	0.98075	0.88473
0.90000	0.99865	0.98684	0.91028

DEPRESS PFK 1 FOR SUMMARY
2 TO REPLOT
3 FOR FURTHER BLOW-UP
31 TO STOP

REPLY AREA

Fig. 2.8.6

```
                              OUTPUT AREA

 :POWER:ALF,4,36,():

  PLOT NO       PARAMETER
     1            0.0500
     2            0.0100
     3            0.0010

    0.60000       0.97498        0.89392        0.65900
    0.60500       0.97591        0.89714        0.66534
    0.61000       0.97682        0.90028        0.67160
    0.61500       0.97770        0.90335        0.67778
    0.62000       0.97855        0.90633        0.68387
    0.62500       0.97938        0.90924        0.68989
    0.63000       0.98018        0.91207        0.69583
    0.63500       0.98096        0.91482        0.70169
    0.64000       0.98171        0.91750        0.70747
    0.64500       0.98244        0.92011        0.71318
    0.65000       0.98315        0.92265        0.71880

DEPRESS PFK  1 FOR SUMMARY
             2 TO REPLOT
             3 FOR FURTHER BLOW-UP
            31 TO STOP

                              REPLY AREA
```

Fig. 2.8.7

Novel features: Plotting of statistical distributions, functions of functions, with multiple-plot facilities for values of a given parameter. Use of a combination of very few precise evaluations, and a cubic spline fitted through eleven points, to produce adequate graphical resolution. Use of blow-up facilities to narrow down the region of interest.

Since this unit was not operational until 2 years after the end of the contract period, its use was restricted to applications in the classroom, in courses on Advanced Scientific Computation and Distribution Theory. This unit was operational and transportable in November, 1977, under MVS on IBM 370/158. The attached figures are photographs from the trial sessions.

CHAPTER 3

THE GMS SYSTEM , ADAPTATION , AND TRANSPORTATION

3.1 Loading from Tape U0120

The partitioned and sequential data sets required to operate the units described in chapters 1 and 2 have been stored on a tape, named U0120, at the Computer Center of the University of Georgia. A copy of this tape, locally labeled U3115, has been sent to the Program Director of Probability and Statistics, at the Office of Naval Research.

Description of Files:

File 1: SYS1.GRAPHLIB

Member Names: #ACTIVE,ANACOV,BASIC,BLOWUP,CALCG, CALCO,COMAP,CURVEFIT,DINTRP,DINTRQ,DUMP,ELLIPSE, FILETST,FORECAST,GENPLT,GWRITERB,GWRITERL,HOPES, LARGE,MAXL,MINGY,MODEL,MULPL,MXCTROL,OMTAB,PATS, PLOTF,PRODFLOW,QUANTAL,RATIO,REGRES,RESCALCB, RES2250,RES2250B,SDES,SPLINE,SPLINEP,SPOOK,SURF, UGUESSIT,WEIBUL,XCTROL

Purpose: These are the load modules of the interactive units; this is the data set which gets modified when units are changed or added. Status of November,1977.

File 2: SYS2.GRAPHLIB

Member Names: QUANBUG,RESCALCB,SHOWFIL

Purpose: Load modules used for experimentation.

File 3: TP.LOAD

Member Names: DEPTS,GETALL,GTR3435,LABELS,LDSAVE, LOGONUTL,MXCTROL,PREPROM,PTPCH,SAMPS2,SESSIONS, TPCHARTS,TSOSMF,TSOSMF2,UADSLIST

File 3(cont'd): Purpose: Utility and monitoring routines.

File 4: SYS1.SSPLIB
Members: IBM Scientific Subroutine Package

File 5: SYS1.EXT2EØA
Random access file; contains displays

File 6: SYS1.W225Ø

FT06F001 and SYSPRINT; for diagnostics (GWRITERB)

File 7: PLOTA
When desired, in a given session, produces image of individual displays, to be processed by CALCOMP plotter.

File 8: PLOTB
Same as File 7, but for permanent storage for two or more sessions.

File 9: SYS1.GMSLIB
Member names: $INITG$,BUILD,EOBINT,FETCH,GBUSY, GFIELD,GPOST,GRINIT,GWAIT,INDEX,INITP,INK,INX,LINGEN, LPINT,PFINT,SCOPLT,STREAL,WAITG,XBLANK,XBLANKS, YENTEST

Utility routines required to add or modify sets on SYS1.GRAPHLIB (see THEMIS Report No. 14 (Penn [18]))

Files 10 , 11 , and 12: Various examples and diagnostics, not needed for use of the system.

Organisation and formats of these data sets are shown in Tables 3.1.1 and 3.1.2 .

TABLE 3.1.1

DATA SETS ON THE TAPE
OF U0120

FILE NO.	NAME OF DATA SET	DSORG	UNIT AT UGA.C.C.	VOLUME AT UGA.C.C.
1	UGA.SYS1.GRAPHLIB	PO	3330	UGD04B
2	UGA.SYS2.GRAPHLIB	PO	3330	UGS004
3	TP.LOAD	PO	2314	UGCPS1
4	SYS1.SSPLIB	PO	3330	UGALB1
5	UGA.SYS1.EXT2EOA	DA	3330	UGS004
6	UGA.SYS1.W2250	PS	2314	UGD04B
7	UGA.PLOTA	PS	3330	UGS004
8	UGA.PLOTB	PS	3330	UGS004
9	UGA.SYS1.GMSLIB	PO	3330	UGD04B
10	JCL TO RUN GMS FOR THE DSN CHANGED	PS		
11	JCL TO RESTORE DATA SET FROM TAPE TO DISK	PS		
12	DOCUMENTATION OF HOW TO USE THE IBM 2250 AND GMS.	PS		
13	ABSTRACT DIRECTION OF THE PROGRAMS IN GMS	PS		

TABLE 3.1.2

DCB FOR EACH DATA SET
OF GMS

NAME OF DATA SET	SPACE USED (TRKS)	RECFM	LRECL	BLKSIZE
UGA.SYS1.GRAPHLIB	350	U	**	7294
UGA.SYS2.GRAPHLIB	22	U	**	2000
TP.LOAD	92	U	**	3625
SYS1.SSPLIB	72	U	**	2000
UGA.SYS1.EXT2EOA	3	F	400	400
UGA.SYS1.W2250	1	FBA	121	121
UGA.PLOTA	1	UBS	364	2000
UGA.PLOTB	2	UBS	364	2000
UGA.SYS1.GMSLIB	21	U	**	7192

Table 3.1.3

JCL for the Loading of Data sets from U0120:

```
//TTTT   JOB(account,password),'LUP',MSGLEVEL=1,CLASS=B,
            REGION=160K,PRTY=9,TIME=10
//*MAIN SYSTEM=SY2,LINES=100
//SE  EXEC  PGM=IEBCOPY
//SYSPRINT  DD  SYSOUT=A
//A DD DSN=TAPE.GRAPH1,UNIT=2400,VOL=SER=U0120,DCB=DEN=2,
//    DISP=OLD,LABEL=(01,NL)
//B DD DSN=UGA.SYS1.GRAPHLIB,UNIT=3330,DISP=(,CATLG),
//    SPACE=(CYL,(50,1,20))
//SYSUT3  DD  UNIT=3330,SPACE=(CYL,(1,1))
//SYSIN  DD  *
XXXX COPY OUTDD=B,INDD=A
//SA  EXEC  PGM=IEBCOPY
//SYSPRINT  DD  SYSOUT=A
//A  DD  DSN=TAPE.GRAPH2,UNIT=2400,VOL=SER=U0120,DCB=DEN=2,
//    DISP=(OLD,PASS),LABEL=(02,NL)
//C  DD  DSN=UGA.SYS2.GRAPHLIB,UNIT=3330,DISP=(,CATLG),
//      SPACE=(CYL,(1,1,1))
//SYSUT3  DD  UNIT=3330,SPACE=(CYL,(1,1))
//SYSIN  DD  *
XXXX COPY OUTDD=C,INDD=A
//SB  EXEC  PGM=IEBCOPY
//SYSPRINT  DD  SYSOUT=A
//A  DD  DSN=TAPE.TPLO,UNIT=2400,VOL=SER=U0120,DCB=DEN=2,
//      DISP=(OLD,PASS),LABEL=(03,NL)
//D  DD  DSN=UGA.TP.LOAD,UNIT=3330,DISP=(,CATLG),
//      SPACE=(CYL,(1,1,4))
//SYSUT3  DD  UNIT=3330,SPACE=(CYL,(1,1))
//SYSIN  DD  *
XXXX COPY OUTDD=D,INDD=A
//SD  EXEC  PGM=IEBCOPY
//SYSPRINT  DD  SYSOUT=A
//A  DD  DSN=TAPE.SSPL,UNIT=2400,VOL=SER=U0120,DCB=DEN=2,
//      DISP=(OLD,PASS),LABEL=(04,NL)
//F  DD  DSN=UGA.SYS1.SSPLIB,UNIT=3330,DISP=(,CATLG),
//      SPACE=(CYL,(1,1,90))
//SYSUT3  DD  UNIT=3330,SPACE=(CYL,(1,1))
//SYSIN  DD  *
XXXX COPY OUTDD=F,INDD=A
//SJ  EXEC  PGM=IEBGENER,REGION=450K
//SYSPRINT  DD  SYSOUT=A
//SYSIN  DD  DUMMY
//SYSUT1 DD DSN=TAPE.EXT2,UNIT=2400,VOL=SER=U0120,DISP=(OLD,PASS),
//      LABEL=(05,NL),DCB=(RECFM=F,LRECL=400,BLKSIZE=400,DEN=2)
//SYSUT2 DD DSN=UGA.SYS1.EXT2EØA,UNIT=3330,DISP=(,CATLG),
//      SPACE=(CYL,(1,1))
//SI  EXEC  PGM=IEBGENER,REGION=450K
//SYSPRINT  DD  SYSOUT=A
```

(JCL for loading of data sets from U0120, cont'd)

```
//SYSIN DD DUMMY
//SYSUT1 DD DSN=TAPE.W2250,UNIT=2400,VOL=SER=U0120,DISP=(OLD,PASS),
//       LABEL=(06,NL),DCB=(RECFM=F,LRECL=121,BLKSIZE=121,DEN=2)
//SYSUT2 DD DSN=UGA.SYS1.W2250,UNIT=3330,DISP=(,CATLG),
//       SPACE=(CYL,(1,1))
//SG  EXEC  PGM=IEBGENER,REGION=450K
//SYSPRINT  DD  SYSOUT=A
//SYSIN  DD  DUMMY
//SYSUT1  DD  DSN=TAPE.PLA,UNIT=2400,VOL=SER=U0120,DISP=(OLD,PASS),
//       LABEL=(07,NL),DCB=(RECFM=U,BLKSIZE=2000,LRECL=2000,DEN=2)
//SYSUT2 DD DSN=UGA.PLOTA,UNIT=3330,DISP=(,CATLG),SPACE=(CYL,(1,1))
//SH  EXEC  PGM=IEBGENER,REGION=450K
//SYSPRINT  DD  SYSOUT=A
//SYSIN  DD  DUMMY
//SYSUT1 DD DSN=TAPE.PLB,UNIT=2400,VOL=SER=U0120,DISP=(OLD,PASS),
//       LABEL=(08,NL),DCB=(RECFM=U,BLKSIZE=2000,LRECL=2000,DEN=2)
//SYSUT2 DD DSN=UGA.PLOTB,UNIT=3330,DISP=(,CATLG),
//       SPACE=(CYL,(1,1))
//SQ  EXEC  PGM=IEBCOPY
//SYSPRINT  DD  SYSOUT=A
//A DD DSN=TAPE.GMSLIB,UNIT=2400,VOL=SER=U0120,DCB=DEN=2,
//       DISP=(,PASS),LABEL=(09,NL)
//D  DD  DSN=UGA.SYS1.GMSLIB,UNIT=3330,DISP=(,CATLG),
//       SPACE=(CYL,(1,1,4))
//SYSUT3  DD  UNIT=3330,SPACE=(CYL,(1,1))
//SYSIN  DD  *
XXXX COPY OUTDD=D,INDD=A
/*
```

Table 3.1.4 JCL for Execution

JCL TO RUN GMS ON UNIV. OF GEORGIA (MVS) NOV. 1977

```
//      A JOB CARD
//*MAIN SYSTEM=SY2,LINES=20
// EXEC PGM=MXCTROL
//GO.STEPLIB DD DSN=SYS1.PL1LIB,DISP=SHR   *
//      DD  DISP=SHR,DSN=TP.LOAD
//      DD DSN=UGA.SYS1.GRAPHLIB,DISP=SHR
//      DD  DISP=SHR,DSN=UGA.SYS2.GRAPHLIB
//GO.FT11F001 DD DISP=(,DELETE),
//      DCB=(,BLKSIZE=0164,LRECL=0160,RECFM=VBS),
//      UNIT=NOSHR,SPACE=(CYL,(1,1))
//GO.FT12F001 DD DISP=(,DELETE),
//      DCB=(,BLKSIZE=0804,LRECL=0800,RECFM=VBS),
//      UNIT=NOSHR,SPACE=(CYL,(1,1))
//GO.FT14F001 DD DUMMY
//GO.FT60F001 DD DUMMY
//GO.FT06F001 DD DSN=UGA.SYS1.W2250,DISP=SHR
//GO.SYSPRINT DD DSN=UGA.SYS1.W2250,DISP=SHR
//GO.WORKDD DD UNIT=3330,SPACE=(CYL,(2),,CONTIG),
//     DCB=(DSORG=DA,RECFM=F,LRECL=3330,BLKSIZE=3330)
//GO.FT05F001 DD DSN=UGA.SYS1.R2250,DISP=SHR
//GO.FT04F001 DD DUMMY,
//     DCB=(BUFNO=1,RECFM=VB,LRECL=84,BLKSIZE=848)
//GO.FT20F001 DD UNIT=NOSHR,SPACE=(CYL,1),
//      DCB=(DSORG=PS,LRECL=32,RECFM=FA)
```

* PL/1 Library

3.1.4 (continued)

```
//GO.FT18F001 DD UNIT=NOSHR,SPACE=(088,(90),,CONTIG,ROUND),
//       DCB=(DSORG=DA,RECFM=VBS)
//GO.FT19F001 DD UNIT=NOSHR,SPACE=(208,(150),,CONTIG,ROUND),
//       DCB=(DSORG=DA,RECFM=VBS)
//GO.FT26F001 DD UNIT=NOSHR,SPACE=(CYL,(1,1),,CONTIG),
//       DCB=(DSORG=DA,RECFM=VBS)
//GO.FT27F001 DD UNIT=NOSHR,SPACE=(CYL,(1,1),,CONTIG),
//       DCB=(DSORG=DA,RECFM=VBS)
//GO.FT28F001 DD UNIT=NOSHR,SPACE=(CYL,(1,1),,CONTIG),
//       DCB=(DSORG=DA,RECFM=VBS)
//GO.FT21F001 DD UNIT=NOSHR,DISP=(,DELETE),
//       SPACE=(CYL,(1,1),,CONTIG),
//       DCB=(,BLKSIZE=0133,LRECL=0126,RECFM=F)
//GO.FT22F001 DD UNIT=NOSHR,DISP=(,DELETE),
//       SPACE=(CYL,(1,1),,CONTIG),
//       DCB=(,BLKSIZE=0133,LRECL=0126,RECFM=F)
//GO.FT16F001 DD UNIT=3330,SPACE=(088,(2000),,CONTIG,ROUND),
//       DCB=(DSORG=DA,RECFM=VBS)
//GO.FT15F001 DD UNIT=3330,DISP=(,DELETE),
//       SPACE=(CYL,(1,1),,CONTIG)
//GO.FT17F001 DD UNIT=3330,SPACE=(088,(2000),,CONTIG,ROUND),
//       DCB=(DSORG=DA,RECFM=VBS)
//GO.FT31F001 DD UNIT=NOSHR,DISP=(,DELETE),SPACE=(CYL,1),
//       DCB=(RECFM=FB,LRECL=80,BLKSIZE=80)
```

```
//GO.FT30F001 DD UNIT=3330,DISP=(,DELETE),SPACE=(CYL,(1,1)),
//      DCB=(LRECL=200,BLKSIZE=816,RECFM=VBS)
//GO.SYSABEND DD SYSOUT=A
//GO.SNAP DD DUMMY
//GO.GRAPHLIB DD DSN=UGA.SYS1.GRAPHLIB,DISP=SHR
//      DD DISP=SHR,DSN=UGA.SYS2.GRAPHLIB
//GO.SYSUPLOT DD DUMMY
//GO.DISPLAY DD UNIT=2F1
//GO.GRAPH1 DD UNIT=AFF=DISPLAY
//GO.FT38F001 DD DSN=UGA.SYS1.MVAP,DISP=SHR
//GO.FT23F001 DD DSN=UGA.SYS1.EXT2EØA,DISP=SHR
//GO.FT08F001 DD DSN=UGA.PLOTA,DISP=SHR,
//      DCB=(RECFM=VBS,LRECL=364,BLKSIZE=2000)
//GO.SYSUT2 DD DSN=UGA.PLOTB,DISP=SHR,
//      DCB=(RECFM=VBS,LRECL=364,BLKSIZE=2000)
```

3.2. Introduction to the Subroutines of the GMS Library *

UGA.SYS1.GRAPHLIB is a load module Partitioned Data Set (PDS) containing all of the subroutines for the GMS library. This PDS has been unloaded to the first file of the tape volume U0120 using the OS Utility IEBCOPY. This tape is 800 BPI, 9 track and non-labeled. The length, in bytes, and the name of each subroutine is included in Table 3.2.1.

PURPOSE

The library of GMS consists of many mathematical and statistical programs needed to perform statistical analyses from the graphics terminal. All subroutines have been compiled through the FORTRAN and Assembler F language processors. To execute a given unit, the user answers questions, enters his data, observes his output and often has some control over the sequence of tasks performed. Every command is given as an explicit instruction in the display.

METHOD

When you see a display on the IBM 2250 screen beginning: "* GRAPHICS MONITOR SYSTEM * THE GRAPHICS DEVICE CAN BE USED BY---,", it means that GMS has started operation. Then you follow the instruction on the display, and depress any one of the lighted PFK's.

* From Yen, Ming, "Transportability of interactive programs", unpublished MS thesis, 1977.

TABLE 3.2.1

SUBROUTINE LENGTH FOR GMS LIBRARY ROUTINES

NAMES	LENGTH IN HEXADECIMAL	LENGTH IN DECIMAL	NAMES	LENGTH IN HEXADECIMAL	LENGTH IN DECIMAL
#ACTIVE	718	1816	MODEL	F6D0	63180
ANACOV	2EAA8	191144	OMTAB	31860	202848
BASIC	E60	3680	PATS	25BA0	154528
BLOWUP	1B848	112712	PLOTF	13800	79872
CALCG	10148	65864	PRODFLOW	29648	169544
CALCO	138F8	80120	QUANTAL	1AB68	109416
COMAP	AD8	2776	RATIO	EE28	60968
GWRITERB	25B8	163952	REGRES	C2B8	49848
FORECAST	28070	84328	SPLINE	F2E0	62176
ELLIPSE	14968	5704	SPOOK	2B9C0	178642
GWRITERL	1648	1648	WEIBUL	6F8	1784
LARGE	670		XCTROL	BC0	3008

The user should simply follow the instruction shown on the screen. He begins by pressing any of the Programmed Function Keys (PFK) except PFK #0. Next he either types $NAMES to see a list of the programs available, or he types $LINK prog, where prog is the name of the program he wants to use. In either case the typed response is followed by an EOB signal, i.e., depressing both the ALT key and the 5 key together. If the user types $NAMES and sees a list of program names, he may continue with $LINK prog as described later.

The user may be asked by a program to respond in any of 3 ways:

1. Type a response followed by an EOB. Any response entered by the typewriter is followed by the EOB sequence.
2. Press the appropriate PFK key on the panel as directed by the program. Some units make this easier by lighting the keys which have special functions at any given state.
3. Place the tip of the light pen on the screen at the appropriate lighted area and operate the foot pedal.

Some features common to most units are:

1. PFK 31 is the "panic button", it terminates most programs. Press PFK 31, when you want to terminate or you don't know what else to do.
2. Even if you are familiar with the program, do not

answer questions before they are asked. If you accidently do this, it may be necessary to press key 31 to terminate and then relink your program. The responses are queued and mismatch may occur if a response is given out of sequence.

3. You can return to the beginning of many programs by pressing key 30.

DESCRIPTION OF PACKAGE PROGRAM

We can divide all the subroutines into the following four parts:

1. Initialization.
2. Graphics Calculator Mode.
3. Statistical Conversational Units for
   - Exploratory Research
   - Data Analysis
   - Student Program Checkout
4. Graphics Utilities:
   - GWRITERB
   - GWRITERL

Table 3.2.2 is included below because it contains an easy-reference summary of the subroutines in the GMS library.

1. INTIALIZATION

MXCTROL see Ref. [18]

MXCTROL constitutes the Master Control Task for the GMS. A mini-monitor system, MXCTROL is the first module

TABLE 3.2.2

PROGRAMS SUMMARY OF SUBROUTINES IN GMS LIBRARY

STEP	SUBROUTINE NAME	PURPOSE	EFFECT OF PFK KEYS
I. INITIALIZATION	MXCTROL	Constitutes the master control task for GMS. MXCTROL is the first module loaded and performs the following steps: · Write-to-operator-with-reply · Attach XCTROL · Enter 'wait' state until the operator types a reply or the sub-task terminates.	None
	XCTROL	Receives control from MXCTROL via the ATTACH macro, initializes and writes first display on the screen. Processes $LINK, $NAMES, $RESET, $END commands.	#0 thru #31
	BASIC	Produce basic character size for COMAP.	None
	LARGE	Produce basic character size for COMAP.	None
	COMAP	Enable the Assembler language programmer to utilize the IBM 2250 graphics terminal as a conversational tool.	None

TABLE 3.2.2 (continued)

STEP	SUBROUTINE NAME	PURPOSE	EFFECT OF PFK KEYS
II. GRAPHICS CALCULATOR MODE	CALCG	The graphics calculator is a single statement interpreter. Statements are written in FORTRAN-like syntax.	None
	CALCO	Plots as a function of X any expression currently available in the calculator mode.	#0, #4, #5, #6, #7, #8, #9, #31, #20, #10, #30
	PLOTF	Plots as a function of X any expression currently available in calculator mode too.	Same as CALCO
III. STATISTICS CONVERSATIONAL UNIT	ANACOV	Analysis of covariance on one response variable and up to 9 concomitant variables. One or two grouping factors with up to 12 levels each.	#0, #15, #30, #31.
	ELLIPSE	Displays clusters by projecting them on various two dimensional subspaces.	#0, #31.
	FORECAST	Second part of Leontieff input-output analysis. Obtains raw material requirements from projected variations and correlations of demand. Network built by "MODEL".	#0, #29, #30, #31.

TABLE 3.2.2 (continued)

STEP	SUBROUTINE NAME	PUSPOSE	EFFECT OF PFK KEYS
III. STATISTICAL CONVERSATIONAL UNIT	RATIO	Investigates the accuracy of empirical estimators of the parameters P and K from the negative binomial distribution.	None
	REGRES	Computes and displays estimated least squares regression line for up to 40 data points.	
	SPLINE	Plot a spline function thru 3 to 50 points. Can be smoothed by use of light pen.	#1, #3 light pen
	SPOOK	Multivariate analysis of variance of irregular data, with up to 10 response variables and 2 grouping factors with up to 12 levels each.	#1, #2, #3, #5, #6.
	WEIBUL	Estimates the parameters of a 2-parameter WEIBUL distribution using order statistics and modified bioassay techniques.	#4, #5, #10, #11.
	PATS	Performs analysis of variance for any complete factorial design of the form $2^R 3^S 4^T < 513$ where $R+S+T < 4$. Obtains response equations and contour maps.	#1, #2, #4, #3, #5, #28, #29, #30, #1, #2, #3.

TABLE 3.2.2 (continued)

STEP	SUBROUTINE NAME	PURPOSE	EFFECT OF PFK KEYS
III. STATISTICAL CONVERSATIONAL UNIT	MODEL	First section of the input-output analysis program. Designed to assist you in constructing a model of the network flow used in forecasting. (see FORECAST)	None
	OMTAB	A column and matrix oriented computing system (OMNITAB-National Bureau of Standards) which is particularly good for matrix calculations Numerical values are stored in a 12 section 80 by 30 worksheet which may be inspected at any time.	None
	PRODFLOW	Product flow analysis, where the user supplies the service properties of a queue. Output may be a statistical distribution of queue lengths and various optimization designs for service.	#20, #1, #31, #5, #25, #26, #27, #28, #29, #30, #10.
	QUANTAL	Analysis of Quantal response data. (PROBIT, LOGIT, WEIBULL, ARCSINE, etc.)	#1, #2, #4, #5, #6, #7, #8, #9, #10, #30, #31.

TABLE 3.2.2 (continued)

STEP	SUBROUTINE NAME	PURPOSE	EFFECT OF PFK KEYS
III. (continued)	BLOWUP	Fitting statistical distributions, power functions, with plots of such functions and blowup.	#1, #2, #3.
VI. GRAPHICS UTILITIES	GWRITERB	Takes an input sequential data set, builds a random access data set as an intermediate data set, and displays the contents of the random access data set by page number upon request by the 2250 user. It uses basic sized characters.	#1, #28, #29, #30, #31.
	GWRITERL	Takes an input sequential data set, builds a random access data set, and displays the contents of the random access data set by page number upon request by the 2250 user. It uses large sized characters.	Same as above

loaded and performs the following steps:

- Write-to-operator-with-reply
- Attach XCTROL
- Enter 'wait' state until the operator types a reply or the sub-task terminates

If the sub-task terminates or the operator replies 'RGMS', a DETACH macro is issued and the above steps are repeated. If the operator replies 'XGMS' the MXCTROL terminates itself.

<u>XCTROL</u> see Ref. [18]

XCTROL receives control from MXCTROL via the ATTACH macro, which means that XCTROL is the daughter task and MXCTROL is the mother task in an Multi-Tasking Environment. XCTROL initializes and writes the first display on the 2250 screen. If the 2250 device power is off at the time, the following message will be written on the computer operator's console seven times.

"*IEA000 ADR,I/OERR,**,0200,4000,#@$GMS"

Where ADR is the unit address, in hexadecimal, of the 2250 device, and the last field, #@$GMS is the job name for GMS, both fields being installation-dependent. The above message is generated by the I/O supervisor, and the explanation for the seven occurrences is that XCTROL performs seven initial I/O sequences to the 2250 display regardless of its ON/OFF condition.

XCTROL then goes into a wait state, until any one of the keys on the PFK is depressed. If PFK #0 is depressed

first, XCTROL is terminated with a user completion code of U1000, the effect being that the initial display is removed from the screen.

The audible alarm sounds, and the same display comes back on the screen. In reality XCTROL terminates, returns control to MXCTROL, is detached and reattached by MXCTROL, at which time XCTROL begins its initialization sequence again.

For an attention interrupt resulting from any PFK other than key #0. XCTROL comes out of the 'wait' state and continues with the next display.

The 2250 screen is partitioned into 2 sections--the larger section called the Output Area, and the smaller section called the Reply Area. Control of the 2250 and abovementioned screen format is provided by COMAP, a Conversational Macro Package for the IBM 2250 assembler-language programmer.

The cursor appears in the first position in the Reply Area and indicates where the next character typed from the alphameric keyboard will appear. Both visibly on the screen and invisibly in the buffer to the 2250 the EOB sequence causes an attention interrupt at the 360 or 370.

The 2nd display depicts the following special commands that are acceptable to XCTROL.

$LINK, $NAMES, $RESET, $END

PROCESSING THE $LINK COMMAND

A model command is $LINK name, where 'name' is supplied by the 2250 operator, i.e. $LINK CALCG. 'NAMES' is the first compared with each of the following:

BASIC, COMAP, LARGE, XCTROL, MXCTROL

These names represent protected GMS control modules and cannot be loaded and executed by the 2250 operator. If a match occurs, then the message:

*** NAME IS A PROTECTED 'GMS' PROCESS ***

is displayed on the 2250 screen, followed by

*** READY FOR COMMAND ***

which signals to the 2250 operator that XCTROL is awaiting a command to be typed in from the alphameric keyboard.

Otherwise, the directory of the graphics library is searched for a match. This search is accomplished by the BLDL macro. If no match occurs for this search, then the message

*** NAME NOT FOUND IN GRAPHICS LIBRARY ***

is displayed on the screen, and XCTROL awaits another command from the 2250 operator as described earlier.

If the directory search produced a match return code=0 from BLDL, then the 2250 is released (RLSEG macro is executed) and the selected module (operand of the $LINK command) is loaded-and-executed via the XCTL macro, which means that XCTROL is replaced in memory by the selected load module, economizing on the use of core storage for GMS.

PROCESSING THE $NAMES COMMAND

A model command is $NAMES and the result is a new display depicting all the program names(members) presently in the graphics library.

These names are read and extracted from the directory of the GMS library by XCTROL which displays three names per line on the 2250.

PROCESSING THE $RESET COMMAND

A model command is $RESET and the result is the erasing of the present display on the screen with the next display being the 2nd display generated by XCTROL.

PROCESSING THE $END COMMAND

A model command is $END and the result is the termination of XCTROL.

PROCESSING AN ILLEGAL COMMAND

If the command is unintelligible to XCTROL, the message:

*** ILLEGAL COMMAND ***

is displayed on the screen.

Each message typed in by the 2250 operator is immediately displayed (up to 40 characters), followed by the action as described earlier. For the processing of all accepted commands, with the exception of the $END command, XCTROL concludes its processing by re-displaying the message:

*** READY FOR COMMAND ***

on the screen indicating to the 2250 operator that XCTROL is done with its processing and is expecting a reply at the time.

<u>BASIC</u> see Ref. [18]

BASIC produces basic sized display characters for COMAP.

<u>LARGE</u> see Ref. [18]

LARGE produces large sized display characters for COMAP.

<u>COMAP</u> (Conversational Macro Package) see Ref. [4]

COMAP enables the assembler language programmer to utilize the IBM 2250 graphics device as a conversational tool with a minimum of programming effort. The following services, normally required of programmers utilizing the problem-oriented routines supplied by IBM (GPS), are provided by COMAP.

- Automatic control and formatting of the output display.
- Dynamic memory allocation and queuing of input message from the alphameric keyboard. Asynchronous to program execution.
- Buffer management in the graphics control unit and handling of all attention interrupts from the 2250 device.

The major design criteria of COMAP was to provide a convenient means for the assembler language programmer to utilize the 2250. The following is a summary of the COMAP

TABLE 3.2.3

THE CHARACTERISTICS OF EACH CHARACTER SET SIZE

CHARACTERISTICS	CHARACTER SET BASIC	CHARACTER SET LARGE
• Maximum number of characters per line	74 (1)	49
• Lines in output area	46	29
• Lines in reply area	2	2
• Characters in output area	3404	1421
• Characters in reply area	148	98
• Load module storage estimate	3669 (2)	1642

(1) Subtract 3 for random message processing to allow for the ID (2 digits plus a blank)

(2) This storage estimate reflects the size of the graphics display buffer maintained by COMAP (including 62 bytes for a linkage table)

macro instructions:

```
symbol  BKSPG  [NC=dd]
symbol  CTRLG  PFK=([pfkmask,pfkaddr]),
               [LP=lpaddr][,ALARM=YES/no]
symbol  DPLYG  NC=dd,BA=bfraddr[,ID=IADDR]
symbol  ERASE  [NL=dd]
symbol  INITG  [SIZE=BASIC/LARGE]
symbol  RLSEG
symbol  RPLYG  NC=dd,BA=bfaddr[,ID=idaddr]
```

2. Graphic Calculator Mode

CALCG see Ref. [18]

CALCG is a single statement interpretive processor, the statements are written in FORTRAN-like syntax with the following rules:

a) Statements are scanned left to right and the 5 arithmetic operations as follows:

"+" for addition

"-" for subtraction

"*" for multiplication

"/" for division and

"**" for power

Assume no pre-determined hierarchy in evaluating the statement. In other words, each statement is scanned (left to right) as a single expression and evaluated as if it were a "polish stack" in memory.

EXAMPLES: Y=A+B*2 means Y=(A+B)*2

Y=A+B*2+C*3 means Y=(((A+B)*2)+C)*3

b) The three allowable terms in an expression are the following:

1) Constant (integer, fixed point, or floating point format)

2) Symbol (8 or fewer characters in each variable name)

3) Function (The available functions will be displayed on the screen upon request by the 2250 operator, as described later in this section.)

c) Only one term is allowed within a single set of parentheses. In other words, CALCG cannot accept multi-termed expressions. Within a variable number of arguments is an operand sub-list appearing in functional expression as long as each item of the sub-list (i.e. the symbols or constants separated by commas) does not violate the above rules concerning nested sets of parentheses, and as long as the number of items in the sub-list does not exceed 5.

Special CALCG commands include:

1. *RESET: To erase all symbols and their values in memory and start CALCG back at the beginning.
2. *OPERATIONS: To list the names of arithmatic operators.
3. *FUNCTIONS: To list the functions provided by CALCG.
4. *LIST: To list each symbol and its current value.

5. *END: To terminate CALCG and to return control back to the Graphics Control Program.

Functions currently available include: LOG, SQRT, TAN, ATAN, SINH, ERF, LGAMA, DLGGM, YORMZ, GAMZ2, CHIZ2, TTX2, BETAZ3, FFZ3, LOG10, SIN, ARSIN, ATANZ, COSH, ERFC, MAX, YORMS, GAMX2, BETAX3, FFX3, CAMNC3, EXP, COS,ARCOS, COSTAN, TANH, GAMMA, MIN, YORMP, GAMP2, CHIP2, TTP2, BETAP3, FFP3, and BETNC5.

The statistical function names above use the following conventions:

(i) Z indicates the ordinate (probability density function)

(ii) X indicates the integral (cumulative distributive function)

(iii) P indicates the percentage points (an X value such that the area to X equals the input value)

(iv) The number at the end of the name is a reminder of the number of input arguments required.

(v) YORM=NORMAL, TT=Student's t, FF=F, CHI=Chi-squre, GAM=Gamma, BETA=Beta, GAMNC3=Non-central Gamma, BETNC5=Non-central Beta (2 types, the confluent hypergeometric has last argument=1, the ${}_2F_1$ hypergeometric has last argument=2, for non-central distribution of r^2, the squre of multiple correlation. Thus BETNC5 $(r^2,q,n,\rho^2,2)$ evaluates the c.d.f. of r^2, with q predictors, sample size N, and non-central value ρ^2.

CALCO

CALCO plots as a function of x any expression currently available in the calculator mode.

EXAMPLE: Y=TP2(X,10)

The user should note that PFK #20 rather than PFK #30 is used to restart this program. PFK #0 rather than PFK #31 terminates the program.

An expression may be plotted for several values of a parameter, say p.

EXAMPLE: Y=TTP2(X,P) where P=10 to 110 by 25

Under the multiple plot option the user should remember that the function to be plotted must have at least two parameters.

After all information about the function has been entered the user is instructed to press PFK #30 for a vector plot or PFK #31 for a point plot. Under the multiple plot option, no matter which key he presses, one plot will be displayed as a vector plot. In this plot the variable parameter is at its lowest value.

After the first plot is displayed, the user may add all other graphs by pressing either PFK #30 or PFK #31. If he presses PFK #10, exactly one plot is added. By repeatedly pressing PFK #10 the user may add more plots to the display.

All plots of the function after the first one will be point plots. Ordinarily a function should not be plotted for more than 7 values of a variable parameter at once.

If too many points on the screen are lighted, the display will be erased and remaining functions plotted separately. To remedy this situation the user might try reducing the interval on the Y axis or he might increase the increment.

CALCO can be used to plot mixtures of distributions where one distribution has a variable parameter.

EXAMPLES: Y=GAMX2(X,6)*1.5+GAMZ2(X,A)*0.4
where A is a variable parameter

Since CALCO, like CALCG, used polish stacking the expression above is equivalent to:

Y=0.6GAMX2(X,6)+0.4GAMZ2(X,A)

PLOTF

PLOTF plots as a function of X any expression currently available in the calculator mode (CALCG).

EXAMPLE: Y=TP2(X,10)
Y=SIN(X)/X
Y=SIN(X)+COS(X)

The option for plotting several values of a parameter is not operational. It is available in another program called BLOWUP.

The unit called BLOWUP permits considerably greater flexibility than either CALCO or PLOTF.

3.3 Adaptation to Non-graphics Equipment

Documentation: Appendix U, appendix vol. V, pp. 192-285, and THEMIS report No. 30 (Hayward and Bargmann [38])

The analysis-of-covariance unit (see Section 2.3) was used as an example for adaptation to a different computer (Control Data Cyber) with a standard, non-graphics terminal (Texas Instruments, Silent 700). This choice was prompted by the following considerations:

(a) Instructions and output appearing on the screen were voluminous, thus the user had to be given an option to see selected portions of output, in random order;
(b) the program had been designed for very fast execution, but required substantial amounts of core - hence an elaborate overlay tree is required;

(c) considerable use was made of program function keys, which had to be simulated on the simpler equipment;

(d) the unit could perform most of its functions without the use of lightpen or graphics.

Formally, the subroutines of the COMFORT package (see Ref. [18]) had to be converted from IBM - graphics made to Control Data teletype mode; most names were simply retained, and their functions altered, so that this adaptation will remain valid for other units.

To permit selective display, at the user's request, use was made of random access ("mass storage") mode. Thus, the user merely types the part number of the desired display, and may go back and forth in random order. To a limited extent this procedure had been followed on the graphics system.

Overlay procedures differ considerably from system to system, even for computers from the same manufacturer. The earlier conversion of a similar unit (batch version of MUDAID, see Section 2.4) from IBM 360/65 to IBM 360/20 DOS served as guidance for this problem. Fortunately, the overlay technique on Control Data systems is very simple.

To simulate the program function keys which, in the graphics unit, could be depressed in lieu of typed entry, the user was directed to prefix "program function key" number by the symbol =. Thus, if instead of replying to a request (by the computer) the user wished to depress program function key 30 for re-start, he or she would type =30.

Novel features: Use of random access mode to permit display of selected portions of voluminous output.

Since the Control Data Cyber at the University of Georgia has been accessible statewide, by telephone or remote job entry, since 1973, this unit was used and demonstrated at other installations (West Georgia College, Armstrong College). This adaptation was also used in a master's thesis in demography.

3.4 Considerations for Adaptation to other Graphics Equipment*

As of this writing (December, 1977) there appears to be little standardization of graphics support programs. IBM announced the availability of a graphics terminal, 3350, in the fall of 1978, and the support programs of the 2250 may be suitable. In the meantime, however, it seems desirable to consider, in detail, the adaptation of a portion of our statistical conversational units to inexpensive and widely available graphics terminals.

We chose to compare the IBM 2250 Graphics console with the Tektronix 4006. Within this section, the following topics are covered: (a) The differences in the features of the Graphics subroutines for the IBM and Tektronix terminal, (b) the Tektronix software, and (c) the use of the cursor on both the IBM and Tektronix models. With this background, we can modify the IBM subroutines to be compatible with the Tektronix 4006.

DIFFERENCES BETWEEN THE IBM 2250 AND TEKTRONIX 4006.

To present an accurate description of the software and internal logic of the Tektronix 4006, it would be best to examine the external or physical attributes of the system, after which we can proceed to the problem of dealing with the conversion of the software.

The Tektronix display terminal is a communications link and display device for use with a wide range of computer systems. The unit is completely self-contained. The display, the keyboard, the operating controls, and the electronics circuity to operate the display and to communicate with the computer are conveniently located and contained within the unit. The terminal is designed to be directly or remotely connected to the computer. Thus, computer operations can be directly influenced by the user

---

*From Yen, Ming, "Transportability of interactive programs, "unpublished MS thesis, 1977.

of the terminal. Finally, there are three basic modes of operation which are a part of the display: (1) Alphanumeric (Alpha); (2) Graphic; and (3) Hard copy.

Let us study how this differs from the IBM 2250.

There are three major differences in the features of the IBM 2250 and the Tektronix 4006. The Tektronix has no Program Function Key (PFK), Light Pen, or Display Buffer, all of which are contained on the IBM 2250. Table 3.4.1 provides a quick comparison of the two terminals. Even though the Tektronix is devoid of the above three features, it does not imply that they cannot be implemented.

PROGRAM FUNCTION KEY

The IBM 2250 has a Program Function Keyboard (PFK) located to the left of the main keyboard. The thirty-two keys, numbered from 0 to 31, are used to permit the user to make decisions in the conversational package program. We do not find the PFK's on the Tektronix 4006.

LIGHT PEN

We can use the Light Pen on the IBM 2250 to communicate with the user program. The operation of the light pen is described in Section 3.1. There is no light pen on the

TABLE 3.4.1

TERMINAL COMPARISON-IBM 2250 VS. TEKTRONIX 4006

FEATURE	IBM 2250	TEKTRONIX (4006)
Program function key	YES	NO
Light pen	YES	NO
Graphics mode	YES	YES
Display buffer	YES	NO
Cursor	YES	YES
Alphameric keyboard	YES	YES

Tektronix 4006.

HANDLING OF THE LIGHT PEN AND PFK

There is no convenient way to simulate the "light pen" on the Tektronix 4006 in graphics mode. The best solution is to move up to the Tektronix 4012 graphics terminal, which offers a "cross-hair cursor" for graphics input mode. This feature is more desirable than the light pen, because any X, Y coordinate can be input to the computer via the cross-hair cursor, whereas, with the light pen, only the X, Y coordinates that are lighted can be input to the computer. A way to represent the PFK number is to type some special character or a name followed by a number from 1 to 31 on the keyboard. (In the CDC Cyber application [38] on Texas Instrument terminals, PFK's were initiated by the character=; "press key 10" would be assumed by "=10").

GRAPHICS SUBROUTINE

Most of the programs in the GMS library depend upon the GRAF and the COMFORT graphics subroutines (Table 3.4.2). The majority of the GRAF and COMFORT subroutine would have to be modified for the conversion to the Tektronix 4006. Table 3.4.2 provides a brief description of each graphics subroutine and how to modify it to run with the Tektronix software. We can now begin to get an overall concept of the relationship that the features have to the internal software support of the Tektronix systems. The importance of the GRAF subroutines, though, should not be discussed

TABLE 3.4.2

MODIFICATIONS TO GMS SUBROUTINE REQUIRED TO

UTILIZE THE TEKTRONIX 4006 TERMINAL

PACKAGE NAME	SUBROUTINE NAME	PURPOSE OR REASON	EDIT SUBROUTINE
COMFORT	GBKSP	To backspace the number of lines on the screen	Modify
COMFORT	GERAS	To erase the lines from the screen	Modify
COMFORT	GWAIT	To place the calling program into a 'wait' state	Modify
COMFORT	GPOST	To take the calling program out of 'wait' state	Modify
COMFORT	GCPFK	To represent attention, interrupts from the IBM 2250	Modify
COMFORT	GRINIT	To initial the character size	Dummy
COMFORT	GRRLSE	To release all main memory assign to COMAP	Dummy
COMFORT	GRRPLY	To transfer the message type by the user	Modify
COMFORT	GRDPLY	To display message on the screen	Modify
COMFORT	GCLP	To represent attention or interruption from the IBM 2250	Modify
COMFORT	GCALM	To set audible alarm	Modify
CRAF	DISPLA	To open graphics data control block and define display variable	Dummy

TABLE 3.4.2 (continued)

PACKAGE NAME	SUBROUTINE NAME	PURPOSE OR REASON	EDIT SUBROUTINE
GRAF	CORDCALL	Coordinate unit (is fixed in Tektronix)	Dummy
GRAF	PLACE	To point the light beam	Modify to TPLOT
GRAF	LINE	To draw a straight line with the light beam	Modify to TPLOT
GRAF	POINT	To draw a point at the location indicated in the call	Modify to TPLOT
GRAF	CHAR	To generate the string consisting of character mode	Modify to CHIN
GRAF	WRFMT$	To write CHAR on the scope under FORMAT control	Modify to CHOUT
GRAF	PLOT	To plot the image on the screen	Modify to TPLOT
GRAF	APEND	To apend more display variables	Dummy
GRAF	RESET	To remove orders from DV area	Dummy
GRAF	UNPLOT	To remove the image from the screen	Modify
GRAF	ERASE	To remove all the image from the screen	Modify
GRAF	BLANK	To blank entire screen	Modify
GRAF	DETAIN	To wait for an attention	Modify to 'wait' state
GRAF	DELAY	To check for an attention	Modify to 'wait' state

TABLE 3.4.2 (continued)

PACKAGE NAME	SUBROUTINE NAME	PURPOSE OR REASON	EDIT SUBROUTINE
GRAF	DETEKT	To check for attention	Modify to 'wait' state
GRAF	LPNAME	No light pen	Dummy
GRAF	CUR$$	To set the position of the cursor	Modify to CURSIS
GRAF	RCUR$	To remove the cursor	Modify to CURSIS
GRAF	SCTDM	To read the data that has been typed in from alphanumeric keyboard	Modify to CHIN
GRAF	SCTDV	To update the data has been typed in	Modify to CHIN
GRAF	DVTDM	To place the resulting data in the FT04F001 dummy buffer	Modify
GRAF	BUFRS	To reset FT04F001 pointers	Modify
GRAF	ALARM	To set audible alarm	Modify
GRAF	SIZE	To compute the length of DV	Modify
GRAF	DVDUMP	No display variable	Dummy
GRAF	LIGHTS	No program function keys	Dummy
GRAF	PRTCHR	To simulate the current display on the printer	Modify

TABLE 3.4.3

GRAPHICS SUBROUTINE IN EACH MEMBERS OF

UGA.SYS1.GRAPHLIB

MEMBER NAME	PACKAGE NAME	GRAPHLICS SUBROUTINE
ANACOV	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GCALM, GERAS, GBKSP, GWAIT.
BLOWUP	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GWAIT.
CALCO	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GCALM, GERAS, GBKSP, GWAIT, GPOST.
ELLIPSE	GRAF	LIGHTS, DETEKT, PLOT, DETAIN, DISPLA, CHAR, CORDCALL, LINE, POINT, PLACE, DCORD, UCORD, UNPLOT, PLACE$, POINT$, ERASE, BLANK, RESET, CUR$$, BUFRS, SCTDV, RCUR$.
FORE-CAST	GRAF	SCTDV, BUFRS, LIGHTS, DETKET, PLOT, DETAIN, RESET, RCUR$, ERASE, CUR$$, CHAR, APPEND, DISPLA, UNPLOT, PLACE$, LINS$.
MODEL	GRAF	BLANK, CUR$$, DISPLA, LIGHTS, PLOT, RCUR$, RESET$, APPEND, ERASE, DSCTDV, UNPLOT, ALARM, CHAR, PLACE$.
PATS	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GCALM, GERAS, GWAIT.
PLOTF	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GCALM, GERAS, GBKSP, GWAIT, GPOST.

TABLE 3.4.3 (continued)

NUMBER NAME	PACKAGE NAME	GRAPHICS SUBROUTINE
PROD-FLOW	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GCALM, GERAS, GBKSP, GWAIT.
QUANTAL	GRAF	DISPLA, PLOT, DETAIN, BUFRS, CUR$$, LIGHTS, SCTDV, ERASE, RCUR$, RESET, CHAR, DETEKT, UNPLOT, PLACE$$, BLANK, CORDCALL, LINES, SIZE.
RATIO	GRAF	CUR$$, DETAIN, DISPLA, PLOT, RCUR$, RESET, UNPLOT, BUFRS, DETKET, CHAR, PLACE$.
REGRES	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GCALM, GERAS, GBKSP, GWAIT, GPOST.
SPLINE	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GCALM, GERAS, GBKSP, GWAIT.
SPOOK	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GCALM, GERAS, GBKSP, GWAIT, GPOST.
WEIBUL	COMFORT	GRINIT, GRRLSE, GRDPLY, GRRPLY, GCPFK, GCLP, GCALM, GERAS, GBKSP.

without a more thorough examination.

We know that the GRAF subroutines provide the FORTRAN programmer with the ability to create and to modify displays composed of points, lines, and characters, and in addition, to plot and to erase the displays on the screen of the IBM 2250. The GRAF subroutines also enable the user to engage the light pen to select parts of the display, which thereby permit the user to communicate with the problem program. The use of the PFK's and indicator lights, and the registering of information from the Alphanumeric keyboard are part of GRAF. But because the Tektronix (among others) have no light pen or PFK's, we should need to modify the subroutines that are directed toward those features. These modifications were outlined in Table 3.4.2. The absence of these features and the modifications inherent on the Tektronix system, may raise questions to its overall effectiveness.

A set of software subroutines has been written by Tektronix to facilitate the use of the Tektronix 4006 display terminal. The software includes basic subroutines for each of the following operating modes: (1) Alphanumeric mode, (2) Graphic Display mode and (3) Graphic Input mode. All of the routines* are written in the FORTRAN language, and thus will run on most computer systems. These subroutines include a subroutine named TPLOT, which is part of the graphics display mode in the Tektronix and would replace the PLACE, LINE, POINT and CHAR subroutines in GRAF.

* Instruction Manual, 4006-1 Computer Display Users Manual, Tektronix, Inc.

It should be pointed out here that the coordinate point system is a little different on the two graphics terminals. With the Tektronix routine, we are not able to see if the plotting on the screen is out of range; therefore, some precaution may need to be taken as to the size of the coordinate system that we are working in. The Tektronix software does provide the user with facilities of a cursor that are compatible with IBM's.

CURSOR

The IBM 2250 and the Tektronix 4006 employ a cursor which is a small line beneath the character position to be entered. That is, when a key is depressed, the corresponding character appears in that position on the screen while the cursor advances to the next position. We use the software subroutines CUR$$ and RCUR$ (on the IBM) to handle this cursor element. With a little modification, this can be changed to CURSES, which is the software subroutine of the Tektronix 4006 for the cursor element; now that we have the cursor subroutines established, the use should be properly examined.

On the IBM, the user controls the positions of the cursor through the use of four keys on the console of the IBM 2250: (1) advance, (2) backspace, (3) continue, and (4) jump. The advance key, when depressed, moves the cursor one character position to the right. The backspace key operates in the same manner as the advance key, the exception being that the direction is reversed. Depression of

the continue key in conjunction with either the advance or the backspace keys will cause the cursor to continue moving forward or backward, respectively, until they are released. Finally, the jump key moves the cursor from its present position to the first character position in the next unprotected block.

By contrast, in the Tektronix terminal, the user controls the position of the cursor through the use of the following four keys on the console of the 4006: (1) page, (2) return, (3) line feed, and (4) cntl H (=backspace). The page key is a special function key which is used to erase the display when the key is depressed. When this occurs, the alpha cursor moves to the upper left corner of the screen (known as the "Home position"). The return key causes the alpha cursor to be positioned at the left hand starting point of the present line. The line feed key causes the cursor to move down one line from the present position.

INPUT/OUTPUT

In the IBM 2250 we read the 2250 buffer table (BT), thereby updating BT to include any data that has been typed in from the alphameric keyboard. These data are stored in the FT04F001 dummy buffer from which they can now be read using a formatted read statement. But we cannot find the graphics buffer, buffer table and display variable on the Tektronix system. We can modify the GMS package read or write directly from the graphics terminal.

We do not have to define the display variable in the Tektronix system, so the subroutines DISPLA, APPEND, SIZE, DVDUMP of GRAF will become dummy subroutines. The subroutines SCTDM, SCTDV, and BUFR$ of the GRAF package reads the data that has been typed in from the alphameric keyboard. We can modify them to "CHOUT" and "CURSES" of the subroutines of Tektronix.

CHAPTER 4
SUPPORTING ALGORITHMS AND SOFTWARE

4.1 Statistical Distribution Package

Documentation: Appendix X, appendix vol. VI, pp. 210-241, and THEMIS report No. 36. (Bouver and Bargmann [42])

Little programming effort is required to provide the standard distribution functions (normal, Gamma, Beta, and its derivations) for moderate precision, not too extreme tails, and integer or half-integer values of parameters. In fact, such programs have been prepared, by the use of series or continued-fraction algorithms, on programmable portable calculators.

However, in the case of real-valued non-integer parameters, extreme tails, or for very large or very small degrees-of-freedom, high precision is quite difficult to attain. But such high precision is required, e.g., in the evaluation of distribution functions belonging to the Pearson class, or for the efficient evaluation of bivariate distributions; in the latter case the truncated probability density function is used as a kernel to determine Gaussian points. A very precise evaluation of moments of such truncated distributions is required, and depends upon distribution functions.

As mentioned in Reference [42], we developed a universal statistical distribution package including normal (YORM), Gamma (GAM), Beta (BET), t (TT), chi-square (CHI), and F (FF). These are FORTRAN library functions; those ending in X evaluate probabilities given abscissas as input; those ending in P evaluate abscissas ("percentage points"), given probabilities as input, and those ending in Z evaluate ordinates, given abscissas. All arguments are real-valued.

A study of efficiencies of various modules showed that, for the normal distribution, for all arguments, and all precision, the two continued fractions (Shenton, for $|x|<2.5$, Laplace for $|x|>2.5$) are best. Percentage points are obtained by Newton-Raphson iteration. A precision of 11 relative places can be guaranteed in the Control Data version

(single precision, where floating point arithmetic is precise to about 13 decimals). A double precision version, with 25-place precision, is also available.

In the Gamma distribution, for high-precision evaluation, the infinite series expansion (non-integer parameters) are useful only if the argument is small (less than the mean) and the degrees-of-freedom are small. The continued fraction algorithm spans a considerably wider range but, if the "degrees-of-freedom" exceed 100, none of the direct evaluations are useful. Limiting forms have been studied (direct approach to normality, approach of the cube root to normality (Wilson-Hilferty)) but they can be used for precision only if the degrees-of-freedom are in the range of one million or greater. For the important intermediate range an algorithm was used which expands the integral into Hermite polynomials, after expansion of the exponent around its maximum value. This routine is both fast and precise for degrees-of-freedom 60 or greater. Thus, we possess a Gamma distribution routine which, on Control Data Cyber, has a relative precision of 11 decimal digits. Double precision versions, to 23 decimals, are also available (though not built in as library functions).

The Beta distribution has much the same properties as the Gamma distribution. Continued-fraction expansions are very good if the larger of the two exponents is less than 50; between 50 and 200 a Hermite expansion works best; beyond that, the approach to a Gamma distribution is very good. The Beta distribution routines built in as library functions in the Control Data Cyber, at the University of Georgia, have a precision of at least 11 significant digits. Double precision routines, to 23 significant digits, are also available.

Incidentally, the "complete" Gamma function (actually log $\Gamma(x)$) is available as single precision (12 significant digits precision) and double precision (23 significant digits precision) under the name DLGGM. It uses the Euler-McLaurin sum formula and is very efficient.

4.2 Non-Central Distributions

Documentation: THEMIS report No. 10 (Thomas [14]) and Appendix B, appendix vol. I, pp. 31-49.

The non-central Gamma distribution (modified Bessel function) and two non-central Beta distributions (${}_1F_1$ and ${}_2F_1$ hypergeometric) have been included in our statistical support package. The non-central Gamma distribution function evaluates the probability integral given the abscissa, shape parameter (exponent plus one) and non-centrality parameter. For use as non-central chi-square distribution, each parameter (and the abscissa) needs to be divided by 2. Through storage of terms in the expansion of the (central) incomplete Gamma integral, the evaluation of the non-central Gamma distribution can be done as quickly as that of the central one.

In the non-central Beta distribution, two cases have been considered (a) a convolution of Poisson terms and central Beta (non-central F, power of the analysis-of-variance test, confluent hypergeometric), and (b) a convolution of negative binomial terms and central Beta (non-central distribution of the square of the multiple correlation, ${}_2F_1$ hypergeometric). When the modules were tested, evaluation was done by two different expansions - the usual infinite series and a finite double sum. As is demonstrated in Ref. [14], the results agreed to 10 significant digits. For the non-central F distribution, the calling sequence is

$$Y = \text{BNC5}\ (x,\ m/2,\ n/2,\ \gamma^2/2,\ 1)$$

where $x = mF/(mF + n)$, the usual conversion of F (with m and n degrees of freedom) into Beta; γ^2 is the non-centrality parameter associated with the non-central χ^2 distribution with m degrees of freedom.

Because of its restricted use for the distribution of multiple correlations, the arguments of the non-central Beta distribution of the second kind have a different interpretation. The calling statement

$$P = \text{BNC5}\ (x,\ q,\ n,\ \rho^2,\ 2)$$

returns in P the P $(R^2 < x)$ (R^2 = square of multiple correlation) given q predictors, sample size n (degrees-of-freedom for error plus 1, if sample is not from a single population), and a population value ρ^2 for the multiple correlation, squared.

The non-central distribution functions are built-in library functions in our Control Data Cyber, and accessible to the IBM graphics and batch systems (see its use in Section 2.8).

4.3 The Pearsonian System

Documentation: Appendix W, appendix vol. VI, pp. 145-209, THEMIS reports No. 28 (Bouver [36]), No. 32 (Bouver and Bargmann [39], and No. 36 (Bouver and Bargmann [42])

Traditionally, the Pearsonian system of distributions has been used to fit empirical data by a member of this extensive class of distributions. However, as early as 1949, percentage points of the Pearsonian distributions were published in the Biometrika Tables, later improved in the CRC Handbook of Probability and Statistics. These tables represent excellent tools for approximating any distribution on the basis of 4 moments. It is a natural extension of the "central limit" normal distribution which is a two-moment approximation and, itself, a member of the Pearson class. Beta and Gamma distributions are special cases, too.

Our concern was to build library functions which enable a user to look up probabilities or percentage points of any member of the Pearsonian class; we also extended the tables far beyond the values given in the Biometrika Table 42, and supplied six place precision.

With one exception (Type IV) the distributions are reducible to the standard distributions described in Sections 4.1. For the type IV evaluation, numerical quadrature techniques were employed which have been developed and described in THEMIS Report No. 26 (Bouver and Lether [24]).

This distribution package is a library function in the Control Data Cyber system, at the University of Georgia. The user may call

$$P = \text{PEARS}\ (x,\ \beta_1,\ \beta_2,\ 2)$$

to obtain $P[Y < x]$ for a Pearsonian distribution with skewness β_1 and kurtosis β_2 (in the notation of the Biometrika Table 42), where Y is standardized (mean 0, variance 1); or he may call

$$Y = \text{PEARS}\ (P,\ \beta_1,\ \beta_2,\ 2)$$

to obtain the percentage point corresponding to the probability P (see instructions to the user on page 148-149 of Appendix volume VI).

Studies are now in progress to obtain bivariate generalizations which are well known theoretically, but quite difficult to implement on computers.

4.4 Convolution of Truncated Gamma

Documentation: Appendix L, appendix vol. III, pp. 216-324, and THEMIS Report No. 21 (Bouver and Bargmann [28]).

The distribution of a sum of independent Gamma variables is itself a Gamma distribution. However, when these distributions are truncated (especially common in applications of the gamma distribution) these sums have rather complicated distribution functions.

A method of characteristic functions was applied to obtain exact distributions in this instance, and tables and graphs were published. In a dissertation by Lavender (1966), the Pearsonian system had been used to approximate the exact distributions. We found that these approximations were very close indeed, in fact, it was the closeness of this agreement which prompted us to extend the library functions of the Pearsonian distributions, as described in Section 4.3.

4.5 Generation of Pseudo-Random Numbers

Documentation: Appendix G, appendix vol. II, pp. 150-245 and THEMIS Report No. 15 (Cannon and Norman [20]).

The purpose of this task was the implementation of a technique proposed by Marsaglia to generate, with great efficiency, pseudo-random numbers under any distribution, with the need for calling an inverse distribution function program occurring only very rarely. Such a package was programmed in IBM Assembler language, and extensively studied for speed of execution and validity in simulating the prescribed distribution.

CHAPTER V

OTHER STATISTICAL TASKS

Most of the tasks described in this section were either begun when the original THEMIS project had a wider definition, or they represented software development which became easy once the interactive packages had been developed and tested.

5.1 Virtual Clustering

Documentation: THEMIS Reports No. 1 and 2 (Bargmann and Graney [1, 2]).

Virtual clustering is the appearance of concentrations of projections of points in n-space on hyperspherical surfaces. The reason for its importance in cluster analysis is the possibility to sort out sample points from mixed multivariate normal distributions. The ordinary ("real") clustering techniques would be of limited value, in this connection, since sample points from multivariate normal distributions are, of course, clustered around the mean, but their projections onto a hyperspherical surface around the mean would be randomly distributed ("random directions"). Departure from such randomness would indicate the presence of more than one mean, i.e., more than one underlying multivariate normal population.

This study resulted in an algorithm and tables of the distribution of distances of k-neighbors. Two approaches are available to the data-analyst: (a) a region may be prescribed (as a fixed opening angle) and the probability that k or more points fall into this region, given random distribution of projections, may be determined; (b) the number of points occurring around a chosen center may be fixed, and the probability determined that the opening angle enclosing these points is Θ or smaller, given that the projections are randomly distributed. Thus, if the observed number of points is larger than the critical value for a given α level, or if the angle is smaller than the critical angle, there would be evidence of significant clustering.

The principles underlying this technique were applied in the interactive unit on cluster configurations (see Section 1.3).

5.2 Stochastic Differential Equations

Reference: THEMIS Report No. 6 (Adomian and Walker [6]), No. 7 (Adomian [7]) and Adomian [22].

The problems of statistical estimation in systems of differential equations have received considerable attention in "compartmental analysis". The usual iterative techniques (commonly called "perturbation" or sensitivity analyses) or expansions into Hermite polynomials would be successful only if the error component is relatively small.

Adomian and some of his doctoral students developed algorithms which could provide solutions even if the random component is very large. In the task studied under this grant, the operator theory was perfected, and some computer simulations were carried out, on our IBM graphics system. Comparison of realizations of simulated results with results obtained by the closure technique showed good agreement. The reason why the simulation unit was removed from the graphics unit later (1974) is that it was quite costly in computer time, thus no photographs are available.

This work is still in progress, even some analytical solutions are becoming available. The most serious drawback is still the high cost of computation even for "closed" solutions, since they require evaluation of second derivative of eigen-vector elements with respect to elements in the model matrix of a system of differential equations.

5.3 Structure and Distance of Logical Patterns

Reference: Appendix A, appendix vol. I, pp. 2-30, and THEMIS Report No. 3, and 8(Patel [3, 11], see also Patel [9]).

A logical pattern in this context is a matrix of order n by p (p possibly much larger than n) in which the entries in each column denote states of certain diagnostic events; (e.g., symptoms of a disease). "Calibration patterns" are stored (patterns of diagnostic events observed at n points in time, when the system is in a known state, e.g., a patient with a known disease) and the same attributes are observed, regularly, to determine whether a given pattern is close to one of the calibration patterns. The n points in time may exhibit dependence (serial correlation).

Since the variance-covariance matrices of such patterns are usually singular, a distance cannot be determined, directly. An assumption is made that the states of the diagnostic events are influenced by the states of just one (or very few) major events, which are artificial. For such a latent-class structure, distances can be calculated.

A simpler version of this technique had been used for the evaluation of intelligence systems. Some application to quality control has been discussed in Reference [3].

5.4 Tools of Analysis for Pattern Recognition

Documentation: Appendix M, appendix vol. IV, pp. 2-62, and THEMIS Report No. 22 (Kundert and Bargmann [29]).

A computer program has been developed which serves as a pre-processor of data, some of which may be ordinal or categorized. Such data are scaled so as to maximize distances between criterion groups (Fisher-Lancaster approach). Multivariate analyses of variance and factor analyses have been performed on data thus pre-processed. These units have had considerable application in food science (to set up scales for subjective attributes and relate them to physico-chemical measures or gas-chromatograms), and in sociological research (effectiveness of birth control information). Examples including frequency-of-repair pattern and real estate valuation are discussed in Reference [29].

5.5 Scaling of Multi-dimensional Categorized Data

Reference: THEMIS Report No. 34 (Chang and Bargmann [40])

The Fisher-Lancaster approach to scaling categorized data is limited to the scaling of two nominal variables against each other (one is usually referred to as a criterion variable). This task attempts to generalize the pre-processing routines described in Section 5.4. The canonical correlation was replaced by a generalized measure (Steel's determinant of a correlation matrix) so that the Lancaster method could be used in the three and more variable case.

The computer program handles up to five nominal variables. Validation studies were performed to demonstrate that the scaling technique does, in fact, calculate correctly the expected values under segments of a multivariate normal distribution.

In 1975 we discovered a union-intersection statistic (eccentricity of an ellipse) which more appropriately generalizes the canonical correlation to three or more sets of variables. Also, this statistic is more easily optimized than the determinant of a correlation matrix. Thus, the pre-processing program has been changed, and the later version is far more efficient.

PUBLICATIONS

[1] Bargmann, R.E. and Graney, R.W. (1969)
"An algorithm for identifying and testing virtual clusters," Dept. of Stat., Univ. of Ga., T.R. No. 42, THEMIS No. 1, 46 pages.

[2] Bargmann, R.E. and Graney, R.W. (1969)
"Tables for significance tests for virtual clusters," Dept. of Stat., Univ. of Ga., T.R. No. 36, THEMIS No. 2, 41 pages.

[3] Patel, H.I. (1969)
"A technique of quality control for logical patterns," Dept. of Stat., Univ. of Ga., T.R. No. 47, THEMIS No. 3, 13 pages.

[4] Penn. L.W. (1969)
"A conversational macro package for the IBM 2250," Dept of Stat., Univ. of Ga., T.R. No. 48, THEMIS No. 4, 35 pages.

[5] Mason, J.D. (1969)
"Some local limit theorems for lattice random variables," Dept. of Stat., Univ. of Ga., T.R. No. 49, THEMIS No. 5, 9 pages.

[6] Adomian, G. and Walker, W. (1970)
"Stationarity theorems for stochastic differential equations," Dept. of Stat., Univ. of Ga., T.R. No. 52, THEMIS No. 6, 51 pages.

[7] Adomian, G. (1970)
"Random operator equations in mathematical physics I," Journ. of Math. Physics, 11, No. 3, pp. 1069-1084 (acknowledgment in following issue).

[8] Patel, H.I. (1970)
"A test of local independence in latent structure analysis," Proc. 1970 Social Statistics, Amer. Stat. Assoc, pp. 342-345.

[9] Patel, H.I. (1970)
"The choice of optimum scores in a Markov chain of order one," Annals Inst. Stat. Math., 22, No. 3, pp. 535-538 (acknowledgment in following issue, p. 135).

[10] Adomian, G. (1970)
"The closure approximation in the hierarchy equations," Dept. of Stat., Univ. of Ga. T.R. No. 53, THEMIS No. 7, 10 pages.

[11] Patel, H.I. (1970)
"Structure and distance of logical patterns," Dept. of Stat., Univ. of Ga., T.R. No. 63, THEMIS No. 8, 114 pages.

[12] Penn. L.W. (1970)
"GPL-1: A graphics programming language for the IBM 2250," Dept. of Stat., Univ. of Ga., T.R. No. 61, THEMIS No. 9, 30 pages.

[13] Mason, J.D. (1971)
"Local theorems in strengthened form for lattice random variables," Annals of Math. Stat., 42, No. 2, pp. 784-788.

[14] Thomas, C.G. (1971)
"Comparison of non-central beta distribution programs," Dept. of Stat., Univ. of Ga., T.R. No. 62, THEMIS No. 10, 77 pages.

[15] Fortson, C.R. (1971)
"Interactive input-output analysis in the presence of variation of demand," Dept. of Stat., Univ. of Ga., T.R. No. 64, THEMIS No. 11, 130 pages.

[16] Knybel, L.J. (1971)
"Queuing analysis for branched processes," Dept. of Stat., Univ. of Ga., T.R. No. 65, THEMIS No. 12, 95 pages.

[17] Schwartz, R.I. (1971)
"A conversational unit for hierarchical discriminant analysis," Dept. of Stat., Univ. of Ga., T.R. No. 66, THEMIS No. 13, 98 pages.

[18] Penn, L.W. (1971)
"An on-line statistical computer system for lay usage," Dept. of Stat., Univ. of Ga., T.R. No. 68, THEMIS No. 14, 2 vols., 400 pages.

[19] Lether, F.G. (1971)
"A generalized product rule for the circle," SIAM Journ. of Num. Anal., 8, No. 2, pp. 249-253.

[20] Cannon, L.E. and Norman, J.E. (1971)
"Pseudo random-number generators for statistical applications," Dept. of Stat., Univ. of Ga., T.R. No. 69, THEMIS No. 15, 151 pages.

[21] Ishee, J.L. (1971)
"Interactive quantal analysis," Dept. of Stat., Univ. of Ga., T.R. No. 74, THEMIS No. 16, 127 pages.

[22] Adomian, G. (1971)
"Linear random operator equations in mathematical physics, II," Journ. of Math. Physics, 12, No. 9, pp. 1944-1948.

[23] Trivedi, S.J. (1971)
"Configuration and classification of clusters in n dimensions," Dept. of Stat., Univ. of Ga., THEMIS No. 17,

[24] Ballengee, A.A. and Bargmann, R.E. (1972)
"An interactive multivariate data analysis program," Dept. of Stat., Univ of Ga., T.R. No. 80, THEMIS No. 18, 150 pages.

[25] Papaioannou, T. (1972)
"A computer-oriented analysis of linear, partially-linear, and qualitative data," Dept. of Stat., Univ. of Ga., T.R. No. 81, THEMIS No. 19, 24 pages.

[26] Norman, J.E. and Cannon, L.E. (1972)
"A computer program for the generation of random numbers from any discrete distribution," Journ. Stat. Comp. Simul., 1, pp. 331-348.

[27] Lether, F.G. (1972)
"A generalized product rule for the triangle," Dept. of Stat., Univ. of Ga. T.R. No. 82, THEMIS No. 21, 18 pages.

[28] Bouver, H. and Bargmann, R.E. (1972)
"On the exact distribution of convolutions of a truncated gamma distribution," Dept. of Stat., Univ. of Ga., T.R. No. 88, THEMIS No. 21, 210 pages.

[29] Kundert, K.R. and Bargmann, R.E. (1972)
"Tools of analysis for pattern recognition," Dept. of Stat., Univ. of Ga., T.R. No. 89, THEMIS No. 22, 250 pages.

[30] Yen, Chang-Wu and Norman, J.E. (1972)
"A conversational unit for fitting data to the Weibull distribution," Dept. of Stat., Univ. of Ga., T.R. No. 90, THEMIS No. 23, 89 pages.

[31] Scott, J.S. and Norman, J.E. (1972)
"A conversational unit for spline function construction," Dept. of Stat., Univ. of Ga., T.R. No. 91, THEMIS No. 24, 73 pages.

[32] Patel, H.I. (1973)
"Quality control methods for multivariate binomial and Poisson distributions," Technometrics, 15, No. 1, pp. 103-112.

[33] Nash, M.S. and Bargmann, R.E. (1973)
"An interactive analysis-of-covariance unit," Dept. of Stat., Univ. of Ga., T.R. 95, THEMIS No. 25, 165 pages.

[34] Bouver, H.S. and Lether, F.G. (1973)
"On the numerical approximation of one, two, and three-dimensional integrals," Dept. of Stat., Univ. of Ga., T.R. No. 98, THEMIS No. 26, 43 pages.

[35] Wood, R.L. (1973)
"A product flow analysis with time-dependence," Dept. of Stat., Univ. of Ga., T.R. No. 99, THEMIS No. 27, 248 pages.

[36] Bouver, H.S. (1973)
"Table of the cumulative standardized Pearson Type IV distribution function," Dept. of Stat., Univ. of Ga., T.R. No. 100, THEMIS No. 28, 19 pages.

[37] Bouver, H.S. (1973)
"Curve fitting by the method of moments," Dept. of Stat., Univ. of Ga., T.R. No. 102, THEMIS No. 29, 90 pages.

[38] Hayward, J.L. and Bargmann, R.E. (1974)
"Methods of conversion of computer-dependent interactive programs - Analysis of Covariance," Dept. of Stat., Univ. of Ga., T.R. No. 105, THEMIS No. 30, 175 pages.

[39] Bouver, H. and Bargmann, R.E. (1974)
"Tables of the standardized percentage points of the Pearson system of curves in terms of β_1 and β_2," Dept. of Stat., Univ. of Ga., T.R. No. 107, THEMIS No. 32, 184 pages.

[40] Chang, C.F. and Bargmann, R.E. (1974)
"Internal multi-dimensional scaling of categorical variables," Dept. of Stat., Univ. of Ga., T.R. No. 108, THEMIS No. 34, 249 pages.

[41] Bingham, S.F. and Bargmann, R.E. (1975)
"An interactive worksheet system for statistical usage," Dept. of Stat., Univ. of Ga., T.R. No. 106, THEMIS No. 31, 291 pages.

[42] Bouver, H. and Bargmann, R.E. (1975)
"Computational algorithms and modules for the evaluation of statistical distribution functions," Dept. of Stat., Univ. of Ga., T.R. No. 110, THEMIS No. 36, 219 pages.

[43] Bond, W.P. and Bargmann, R.E. (1976)
"Graphical aids for statistical computation," Dept. of Stat., Univ. of Ga., T.R. No. 112, THEMIS No. 37, 335 pages.

MAJOR PRESENTATIONS (Invited)

R. E. Bargmann

North Carolina State University, Raleigh, February 1970

CRICISAM 10th Annual Conference, Tallahassee, Fla., March 1970

Lackland Air Force Base, San Antonio, Texas, August 1970

Infantry Agency, US Army Development Command, Ft. Benning, Ga., August 1970

Eglin Air Force Base, Fla., September 1970

Redstone Arsenal, Huntsville, Ala., December 1970

Lecture Series at Lockheed, (Ga. Tech. Program) Spring 1971 and Summer 1971

American Statistical Association, Ames, Iowa, April 1972

Chemical Division, Am. Soc. for Quality Control & Am. Statist. Assoc., Knoxville, Tenn., September 1972

Lecture Series at Medical College of South Carolina, Charleston, S.C., September 1972

American Statistical Association, New York, December 1973

Emory University, Atlanta, Ga., April 1974

Oklahoma State University, Stillwater, Oklahoma, February 1974

Univeristy of Arkansas, Fayetteville, Arkansas, April 1974

West Georgia College, April 1974

8th Symposium on Interface, Los Angeles, February 1975

Armstrong State College, May 1975

American Statistical Association, Atlanta, Ga., August 1975

W. P. Bond

Chemical Div., ASQC, and American Statistical Association, Knoxville, Tenn., September 1972

American Statistical Association, New York, December 1973

S. Bingham

American Statistical Association, New York, December 1973

H. S. Bouver

American Statistical Association, Boston, August 1976

C. F. Kossack

Lackland Air Force Base, San Antonio, Texas, August 1970

Eglin Air Force Base, Fla., December 1970

DEGREES AWARDED

(research supported by this grant)

Ph.D. in Mathematics:

Dr. Winston Walker (1970); major prof.: Adomian

Ph.D. in Statistics:

Dr. Richard W. Graney (1969); major prof.: Bargmann
Present position: Assoc. Prof., Univ. of Wisconsin at Platteville

Dr. H. I. Patel (1970); major prof: Bargmann
Present position: Statistician, CIBA-GEIGY Corp., Summit, N.J.

Dr. Lucius W. Penn (1971); major prof.: Bargmann
Present position: Assist. Prof., Univ. of Georgia

Dr. Surendra J. Trivedi (1971); major prof.: Bargmann
Present position: Assoc. Prof., Mankato St. College, Mankato, Minn.

Dr. Kenneth R. Kundert (1972); major prof.: Bargmann
Present position: Prof. of Mathematics, Univ. of Wisconsin at Platteville

Dr. Chit-Fu Chang (1974); major prof.: Bargmann
Present position: Assoc. Prof., Gardner-Webb College, Boiling Springs, N.C.

Dr. Hubert S. Bouver (1975); major prof.: Bargmann
Present position: Assist. Prof., SUNY at Plattsburgh, N.Y.

Dr. Steven F. Bingham (1975); major prof.: Bargmann
Present position: Statistician, Veterans Administration Hospital, Perry Point, MD

Dr. Walter P. Bond (1976); major prof.: Bargmann
Present position: Manager, System Testing, Computer Science, Corp., Jacksonville, Fla.

Continuation of Degrees Awarded

M.S. in Statistics (with thesis)

Carl R. Fortson (1971)	thesis advisor:	Bargmann
Judy L. Ishee (1971)	" "	Bargmann
Robert I. Schwartz (1971)	" "	Bargmann
Carlton G. Thomas (1971)	" "	Bargmann
Ann A. Ballengee (1972)	" "	Bargmann
Lovick E. Cannon (1972)	" "	Norman
Lois L. Knybel (1972)	" "	Bargmann
Janice S. Scott (1972)	" "	Norman
Chang-Wu Yen (1972)	" "	Norman
M. Elizabeth Nash (1973)	" "	Bargmann
R. L. Wood (1973)	" "	Bargmann
Judy L. Hayward (1974)	" "	Bargmann

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) UNIVERSITY OF GEORGIA ATHENS, GEORGIA	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP UNCLASSIFIED

3. REPORT TITLE

9 FINAL REP[redacted]T., 6 Interactive Statistical Software.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

10 ROLF E. BARGMANN [redacted] HUBERT BOUVER

6. REPORT DATE 11 JAN[redacted]78	7a. TOTAL NO. OF PAGES ~~155~~ 12 168p.	7b. NO. OF REFS 43

8a. CONTRACT OR GRANT NO. 15 N000 14-69-A-0423 b. PROJECT NO. NR 042-261 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 14 THEMIS-UGA-33 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) THEMIS NUMBER 33

10. DISTRIBUTION STATEMENT

REPRODUCTION IN WHOLE OR PART IS PERMITTED FOR ANY PURPOSE OF THE UNITED STATES GOVERNMENT.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY OFFICE OF NAVAL RESEARCH WASHINGTON, D.C.

13. ABSTRACT

FINAL REPORT (NARRATIVE PORTION)

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
AFFINE TRANSFORMATION						
ANALYSIS OF COVARIANCE						
ANALYSIS OF VARIANCE						
APPROXIMATION						
BETA DISTRIBUTION						
BIO-ASSAY						
BIVARIATE NORMAL DISTRIBUTION						
BRANCHED PROCESSES						
CATEGORICAL SCALING						
CHARACTERISTIC FUNCTION						
CLASSIFICATION						
CLUSTER ANALYSIS						
COMPOSITE RULES						
COMPUTER GRAPHICS						
COMPUTER LANGUAGE						
CONFIDENCE INTERVALS						
CONICAL PRODUCT RULES						
CONTINGENCY TABLES						
CONVERSATIONAL COMPUTATION						
CONVOLUTIONS						
CUBATURE						
CUMULATIVE DISTRIBUTION FUNCTION						
CURVE FITTING						
DATA ANALYSIS						
DESIGN OF EXPERIMENTS						
DIGAMMA FUNCTION						
1-2-3 DIMENSIONAL INTEGRALS						
DISCRIMINANT ANALYSIS						
DISTANCE						
ESTIMATION						
FACTOR ANALYSIS						
FORECASTING						
FOURIER TRANSFORMS						
GAUSS-LEGENDRE						
GAUSS-LEGENDRE QUADRATURE						
GAUSS-LOBATTO						
GENERALIZED PRODUCT RULES						
GRADUATION						
GROWTH CURVES						
HIERARCHICAL DISCRIMINATION						
HYPERCUBES						
HYPERSPHERES						
INCOMPLETE BETA DISTRIBUTION						
INCOMPLETE GAMMA DISTRIBUTION						
INPUT-OUTPUT ANALYSIS						
INTERACTIONS						
INTERACTIVE PROGRAMMING						
INTERPOLATION						
INVERSE TRANSFORMS						
LATENT CLASS ANALYSIS						
LEARNING MODELS						
LEAST SQUARES FITTING						
LINEAR MODELS						
LOCAL LIMIT THEOREM						

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
LOCATION						
LOG-LOG TRANSFORMATION						
MANAGEMENT INFORMATION SYSTEM						
MAXIMUM LIKELIHOOD ESTIMATION						
METHOD OF MOMENTS						
METHOD OF TRANSLATION						
MINIMUM X^2 METHOD						
MOMENTS OF TRUNCATED GAMMA						
MONTE CARLO						
MULTIPLE CORRELATION						
MULTIVARIATE ANALYSIS						
MULTIVARIATE POLYNOMIALS						
NEWTON-COTES						
NEWTON-RAPHSON						
NONCENTRAL F						
NONLINEAR ESTIMATION						
NONLINEAR MODELS						
NUMBER-THEORETIC						
NUMERICAL APPROXIMATION						
ON-LINE COMPUTATION						
OPERATIONS RESEARCH						
PARALLEL TANGENTS (PARTAN)						
PARTIALLY LINEAR MODELS						
PATTERN RECOGNITION						
PEARSONIAN DISTRIBUTIONS						
PEARSONIAN SYSTEM						
PEARSON'S CURVE						
PERCENTAGE POINTS						
PROBABILITY LEVEL						
PROBIT ANALYSIS						
PRODUCT FLOW ANALYSIS						
PSEUDO RANDOM NUMBERS						
QUADRATURE						
QUANTAL ANALYSIS						
QUEUEING THEORY						
QUEUES						
REGRESSION						
SCALE PARAMETERS						
SCALING						
SHAPE						
SIMULATION						
SPLINE						
STANDARDIZED MOMENTS						
STATISTICS						
STATISTICAL COMPUTATION						
STATISTICAL DISTRIBUTION PACKAGE						
STOCHASTIC DIFFERENTIAL EQUATIONS						
STOCHASTIC SYSTEMS						
SUM OF TRUNCATED GAMMA DISTRIBUTIONS						
TRANSIENT QUEUE ANALYSIS						
TRIANGLE						
TYPE IV						
TYPE V						
WAVE PROPAGATION IN STOCHASTIC MEDIA						
WEIBULL DISTRIBUTION						

DD FORM 1 NOV 65 1473 (BACK)
S/N 0101-807-6821

UNCLASSIFIED
Security Classification

A-31409